Agriculture Module - Appendix

Collection 11

Version 1

General coordinator
Eliseu Weber

Team
Kênia Samara Mourão Santos
Paulo Domingos Pires Teixeira Junior
Clebson Ismael dos Santos e Silva
Guilherme Vanz dos Santos
Mylena Cavalcante Mourão

August, 2026

Document Overview

This appendix details the methods used for the Collection 10 of products from the MapBiomas' Agriculture module. This document is divided into the following structure:

A. Irrigation Systems
Description of the classification methods for the irrigated agriculture classes included in the "Irrigation Systems" map.

B. Number of Cycles
Description of the methods for the "Number of Cycles" maps. Currently in its beta version.

C. **Land Use Land Cover - Second Season**
Description of the methods for the Second Season map. Currently in its beta version.

Description of the classification methods for all agriculture and forest plantation classes included in MapBiomas' Land Use Land Cover and the "Agricultural Use" map inside the Agriculture module can be found here.

The MapBiomas organization in GitHub has repositories for all the network's initiatives and modules. The 'Agriculture' repository contains the scripts used in all products inside the Agriculture module and is available at:

➔ Agriculture: https://github.com/mapbiomas/brazil-agriculture

A. Irrigation Systems

1 Overview of the classification method

The MapBiomas project produces, among other land use and land cover classes, annual irrigation agriculture maps in Brazil from 1985 to the present. The first irrigation agriculture map from MapBiomas was released on Collection 5, comprising from 2000 to 2019, with maps of center pivot irrigation, covering all Brazil, and other irrigation systems, covering only the semiarid region. In Collection 6, the irrigation rice class was added and the other classes were extended to the 1985-2020 period. In Collection 7.1, in addition to the classes from the previous Collections, a new type of information about irrigation was added, the pivot dynamic. Pivot dynamics consists in presenting individualized characteristics of each pivot, such as number of cycles per year, dates of start and end cycles, and average daily precipitation. In the Collection 8 and Collection 9, the information about center pivot, other irrigation systems and pivot dynamics were reviewed and the 2022 and 2023 years were added, respectively. In Collection 10, the pivot dynamics product was discontinued in favor of a pixel based approach to retrieve the same information, which started with the number of cycles product in the Agriculture module.

2 Center pivot irrigation systems

The first attempts in the MapBiomas project for mapping center pivot irrigation systems came through the Next Generation Mapping (NexGenMap) project. The objective of this initiative was to develop machine learning algorithms, tools and methods for producing the most current, detailed and accurate maps of land use and land cover using daily PlanetScope imagery, cloud computing, and new artificial intelligence algorithms. In the NextGenMap project, artificial intelligence algorithms were developed to map center pivot irrigation systems using PlanetScope imagery in the Cerrado biome (SARAIVA et al., 2020).

In MapBiomas context, the mapping of ‘Center pivot irrigation systems’ was performed using Landsat imagery and an adapted U-Net architecture (RONNEBERGER et al., 2015), an image segmentation convolutional neural network architecture. The adapted U-Net architecture was trained with two different sets of samples, one set with center pivot irrigation systems samples and other with irrigated rice samples. To increase the temporal and spatial consistency of the final maps, the raw result was post-processed using temporal and spatial filters (Figure 1A).

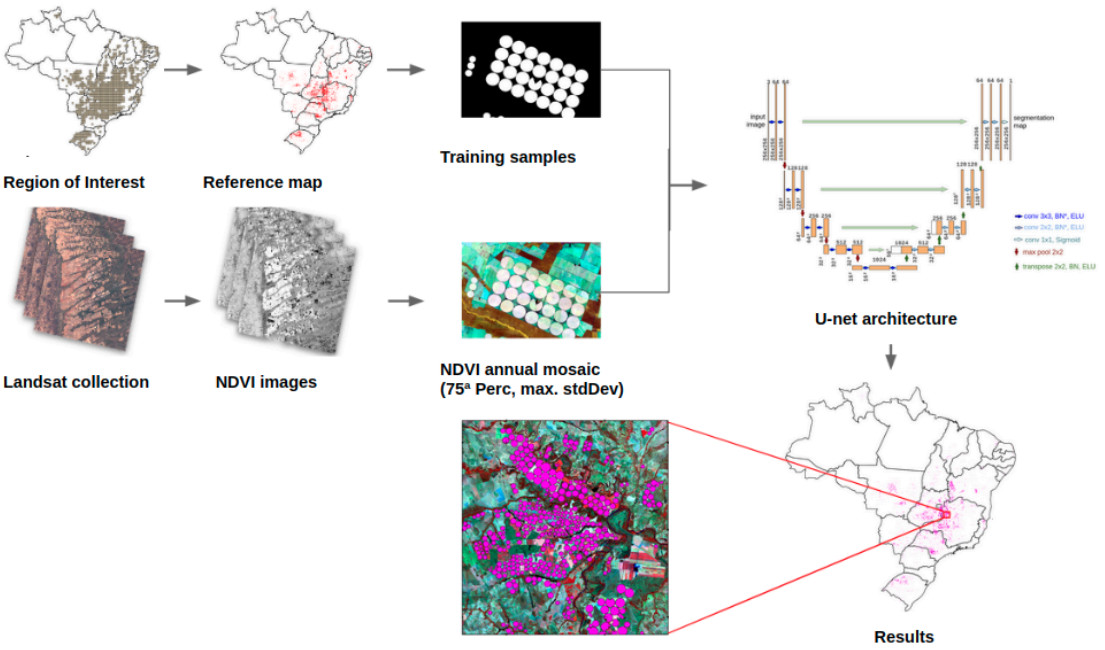


Figure 1A: steps of the mapping process of two center pivot irrigation systems.

2.1 Image selection

The mapping of the center pivot irrigation systems used annual mosaics generated from available images in each year. However, in this Collection, as was only added the 2022 year in the temporal series from the last Collection, the images from the period of 1985 to 2021 were Landsat Collection 1 Tier 1 TOA, and for 2022 were Landsat Collection 2 Tier 1 TOA. In addition, only images with less than 80% cloud cover and shadows were considered.

2.2 Definition of regions for classification

The reference maps used for categorizing center pivot irrigation systems were generated through a collaboration between the Brazilian National Water Agency (ANA) and Embrapa Milho e Sorgo, corresponding to the years 1985, 1990, 2000, 2005, 2010, 2014, and 2017 (ANA, 2019). These mappings were produced based on visual interpretation of imagery acquired from Landsat 5, Landsat 8, and Sentinel 2A/2B satellites, alongside high-resolution images (<1 meter) sourced from Google Earth.

For the delimitation of the study area, the Brazilian territory was divided into blocks of 0.5' x 0.5' degrees (~300 thousand ha each). Only blocks with occurrence of center pivot irrigation systems in any of the reference map years were selected. Figure 2A shows the 723 chosen blocks distributed across an area of approximately 212 million hectares to map center pivot irrigation systems in Brazil.

Reference maps (ANA, 2019a)

Study area used for mapping of center pivot irrigation systems

Figure 2A. Study area for the mapping of center pivot irrigation systems in Brazil.

2.3 Classification

2.3.1 Classification scheme

The classification scheme of the center pivot irrigation considered only two possible classes for each pixel, center pivot irrigation, and non-center pivot irrigation.

2.3.2 Feature space

The feature space created for the center pivot irrigation systems mapping aimed to obtain the characteristics of the pivot at the time they were cultivated, as well as to highlight the differences in relation to the other targets, such as other agriculture areas, pasture, forest formation, etc. Therefore, three metrics were selected that showed the best results to distinguish the pivots in relation to the other targets:

- NDVI_p75, 75th percentile of NDVI values for all images;
- NDVI_p100, 100th percentile, or maximum value, of the NDVI values of all images, and;
- NDVI_stdDev, the standard deviation of the NDVI values for all images.

The mosaic generated is composed by the selected metrics. Each metric corresponds to a band in the image, as shown in Figure 3A.

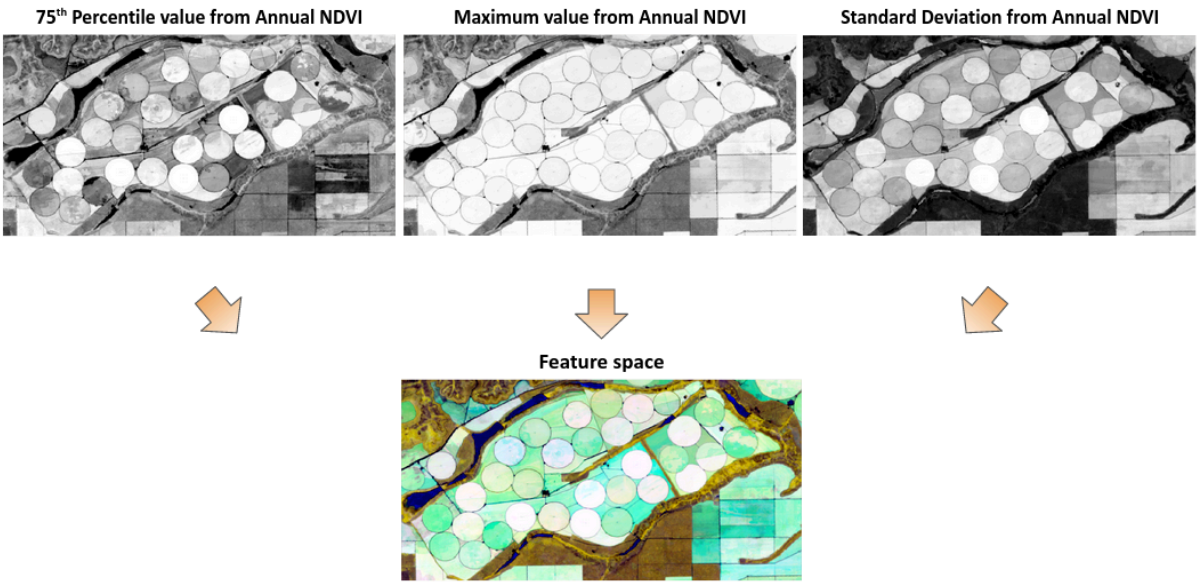


Figure 3A. RGB visualization (NDVI p75, NDVI Maximum, NDVI stdDev) of an image used for training and mapping of the center pivot irrigation systems, generated for the year 2017.

The use of images with only three bands improved the process of training and classifying the pivots since the reduced amount of bands, consequently reduced the computational infrastructure necessary for the processing of this data.

2.3.3 Classification algorithm, training samples and parameters

Due to the extensive study area (~212 Mha) and computational limitations, the model was trained using only a subset of blocks chosen from the population of 723 blocks.

The choice of sample data is an important step for training Deep Learning models, since the samples must represent all the spatial and spectral variability of the population. For this, stratified sampling was performed based on the pivot area obtained from the reference maps. The sampling considered three strata: with low, medium and high coverage of center pivot irrigation systems. The stratum containing blocks with low coverage was created from the blocks whose pivot area was less than or equal to the median of the area of all blocks, that is, 50% of the blocks (361 blocks). The stratum with the high coverage was created from blocks whose sum of the area of its pivots covers about 50% of the pivot area of the entire population (total of 41 blocks). Finally, the remaining blocks (321 blocks) were used to create the layer with blocks containing a medium cover of center pivot irrigation systems. After creating the stratum, 20 blocks were randomly chosen for training and 10 blocks for testing in each of the three stratum. The training blocks were used to calibrate the model, while the test blocks were used later for the accuracy analysis of the model. Figure 4A illustrates the spatial distribution of the stratum and blocks chosen for training and testing the population model.

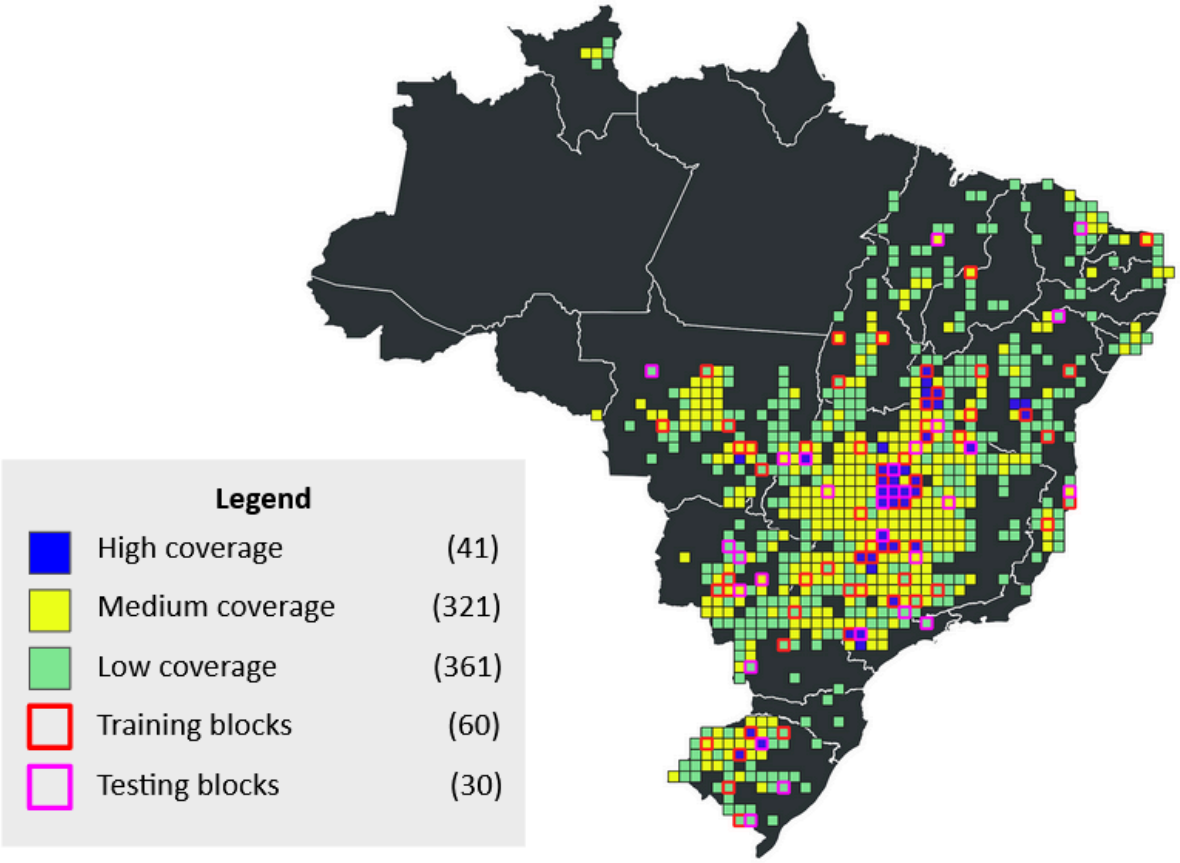


Figure 4A. Spatial distribution of high, medium and low center pivot irrigation cover stratum and location of the blocks used for training and testing the model in Brazil.

As mentioned earlier, an adaptation of the U-Net convolutional neural network architecture was performed to map the center pivot irrigation systems. Figure 5A illustrates the modified U-Net architecture created.

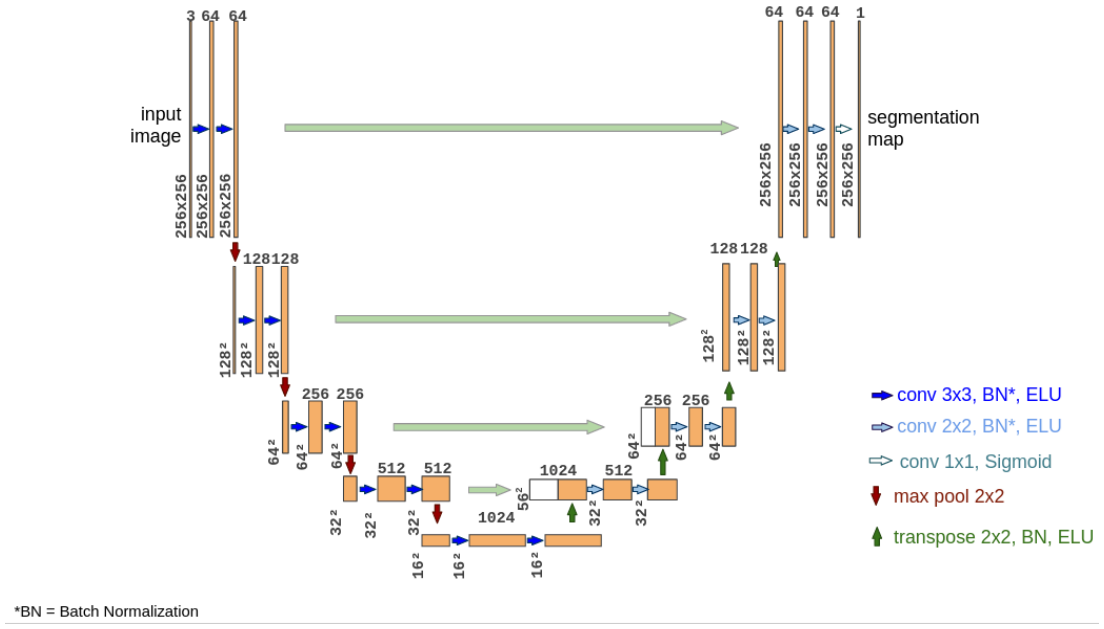


Figure 5A. Adapted U-Net architecture, with its layers and connections, used for the mapping of center pivot irrigation systems.

This architecture was developed in Python, using the TensorFlow 2.0 library. The entire training and mapping process was carried out using the Google Colab platform using Google Drive to access the annual mosaics (generated in Google Earth Engine). Table 1A presents some hyperparameters used during model training.

Table 1A. Hyperparameters for training the modified U-Net architecture.

Hyperparameter	Value
Chip size	256 x 256 pixels
Batch size	20
Epochs	100
Learning rate	0.001

The 2017 reference map was used for model training. In the training blocks, chips with 256 x 256 pixels were generated, 75% were allocated to the training data set and 25% for the validation data set. Figure 6A illustrates the process of subdividing training blocks into smaller chips to be used as input for model training.

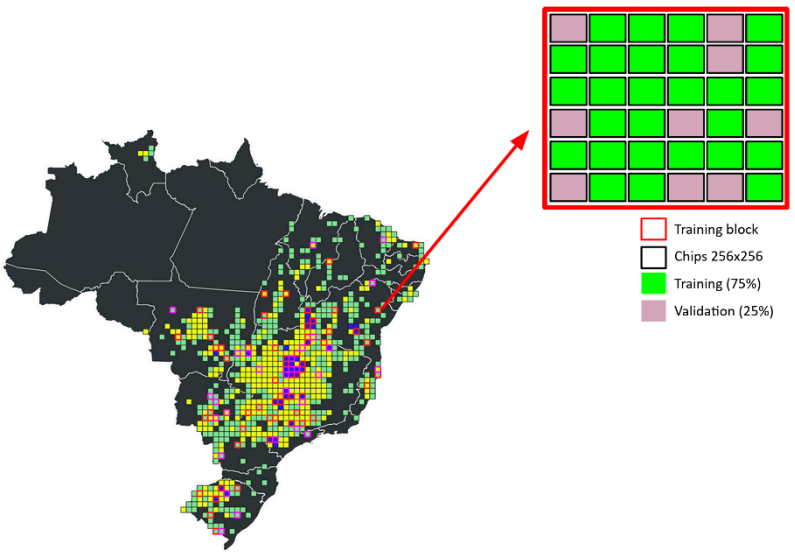


Figure 6A. Examples of the training and validation chips allocated within the training block of the model.

The training set was used to learn the model and the validation set used to perform initial validations during model learning.

After the network training process was completed, the classifier was applied throughout the Brazilian territory. In this step, 1024 x 1024 pixel chips were used. Increasing the size of the chips at the time of sorting not only decreases problems generated by the edges of the chips but also increases the memory capacity required for processing. Therefore, it was necessary to decrease the batch size to 1.

2.4 Post-Classification

2.4.1 Temporal filter

The temporal filter employed for center pivot irrigation systems maps consisted of a five-year moving window. Within this window, the targeted pixel was modified according to two guiding rules:

1. the pixel is changed to center pivot if at least one of the two previous years and at least one of the two subsequent years, that pixel was mapped as a pivot, indicating a possible model omission error;

2. pixels that were mapped as pivots only in the assessed pixel of the five-year window, indicating a possible inclusion error, have been removed from the classification.

2.4.2 Spatial filter

In the center pivot irrigation systems mapping it was used a spatial filter based on the erosion operation followed by an expansion operation using a circular kernel with a radius of 60 meters. This spatial filter helped to eliminate noise generated by the mapping, as well as smoothing the edges of the center pivot irrigation (Figure 7A).

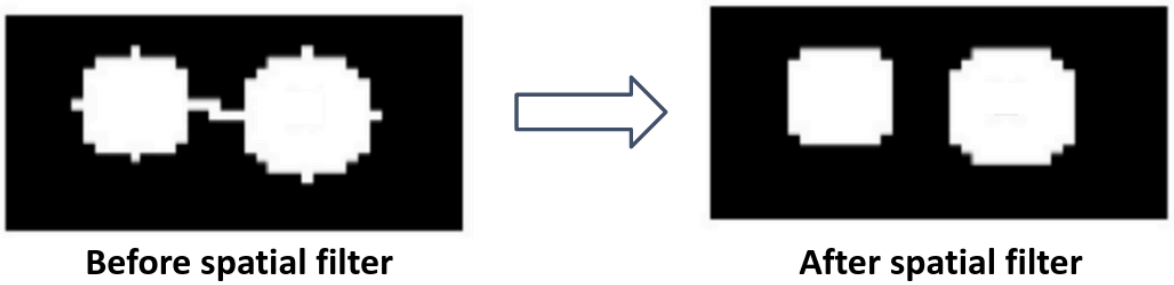


Figure 7A Example of correction of the spatial filter (on the right) in a classification that presents noise on the edges (on the left).

2.5 Validation strategies

The preliminary validation of the center pivot irrigation model used the test blocks of the 2017 mapping (see Figure 4A), as these blocks were not used for the model training. From

the reference map, the user's and producer's accuracy was calculated for each of the individual stratum and also considering all strata at the same time. Table 2A presents the results of the preliminary model validation.

Table 2A. Preliminary validation of the center pivot irrigation mapping for the year 2017, using the test blocks selected in each stratum.

Stratum	Producer's Accuracy	User's Accuracy
Low coverage	40.87%	71.39%
Medium coverage	86.37%	91.62%
High coverage	84.16%	96.19%
All strata	83.97%	95.38%

The preliminary accuracy analysis showed that, in 2017, the model performed better in regions with higher center pivot coverage. Considering all strata in 2017, the model presented an omission error of 16% and an inclusion error of 5%.

2.6 Results

Comparing the area mapping results between MapBiomas Collection 11 and the ANA/EMBRAPA dataset reveals that, until approximately 2010, the area mapped by MapBiomas closely aligns with the values reported in the ANA/EMBRAPA dataset. However, in the subsequent comparison years, a more substantial discrepancy emerges, with MapBiomas Collection 11 mapping smaller areas than ANA/EMBRAPA.

Over the temporal series, the ANA/EMBRAPA dataset indicates an initial area of approximately 0.03 Mha in 1985. This area steadily increases, reaching its peak of 1.90 Mha in 2020. In comparison, MapBiomas Collection 11 starts at approximately 0.08 Mha in 1985 and shows consistent growth, reaching approximately 1.54 Mha in 2020 and 2.07 Mha in 2025 (Figure 8A).

In addition, it is important to point out that these divergences between MapBiomas and ANA/EMBRAPA, especially in the most recent comparison years, underscore the importance of considering their distinct methodologies and data sources. These differences may be partly related to the omission of small center pivots from the MapBiomas mapping.

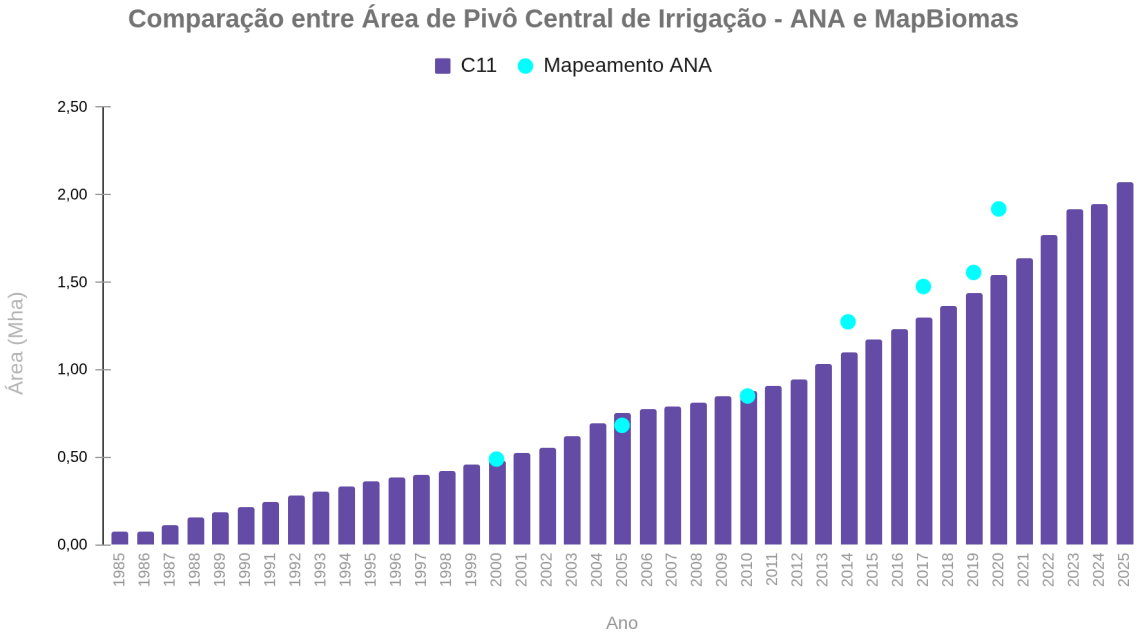


Figure 8A: Results of automatic mapping of center pivot irrigation systems in Brazil based on Landsat images for the period from 1985 to 2025 compared to surveys carried out by the ANA (ANA, 2022).

3 Irrigated rice

Since Collection 9, rice is mapped using Random Forest algorithm to complement the results obtained with U-net U-Net architecture, a segmentation convolutional neural network architecture (RONNEBERGER et al., 2015). Samples are obtained by reference maps from the National Water Agency (ANA, 2021b) and the National Supply Company (ANA e Conab, 2020). To increase the temporal and spatial consistency of the final maps, the raw result is also post-processed using temporal and spatial filters.

The irrigated rice class present in the Irrigation Systems map is the same as the rice class present in MapBiomas' Land Use Land Cover map. Details of the rice mapping methodology can be found in the Agriculture and Forest Plantation Appendix.

4 Irrigated agriculture in semi-arid region

In Collection 9, a notable improvement in the approach employed for classifying the 'Other irrigation systems' class is worth mentioning. The methodology maintains its foundation in pixel-by-pixel mapping through the utilization of the Random Forest algorithm. However,

within Collection 9, a novel reference map of irrigated agriculture in the semiarid region was acquired in collaboration with the National Water Agency (ANA). Additionally, it was also obtained from ANA, a regular grid, demarcating regions within the semiarid region where occurrences of irrigated crops are prevalent. These augmentations to the methodology serve to enhance the accuracy and precision of the mapping process for the 'Other irrigation systems' class. Figure 9A presents the flowchart of the methodology for 'Other irrigation systems' classification

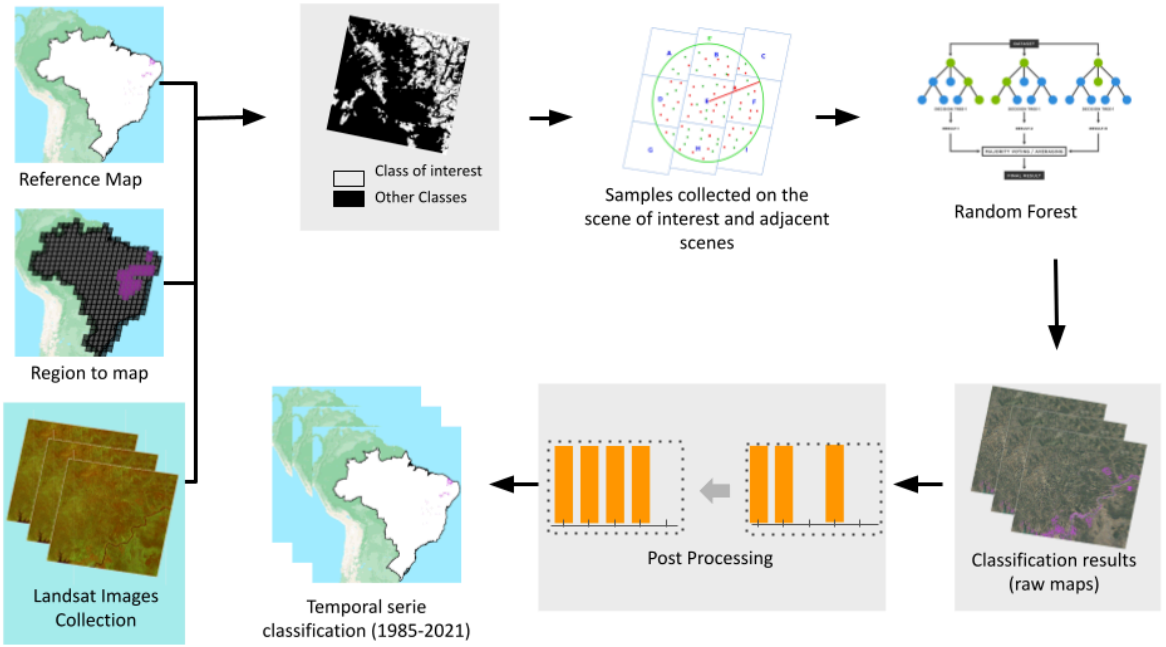


Figure 9A. Classification process for mapping 'Other irrigation systems' in MapBiomas.

4.1 Image selection

In this new approach for Collection 9, the mapping process relied on the use of yearly TOA (Top of Atmosphere) mosaics from Landsat Collection 2. These mosaics were supplemented by a set of spectral indices and statistical measures. This combination aimed to emphasize and distinguish between areas of irrigated agriculture and the native vegetation. By incorporating these spectral indices and statistical measures, the methodology gains the ability to identify and highlight the distinct features that characterize irrigated agricultural areas within the context of the surrounding natural vegetation.

4.2 Definition of regions for classification

In the mapping of other irrigation systems, the study area was restricted to the Brazilian semi-arid region. In this region, due to water requirements, irrigation is almost a mandatory requirement to reduce production risks and/or increase productivity.

In the previous Collections, a total of 34 municipalities were employed as the mapping area (selected due to their substantial expanse of irrigated agriculture). However, several of these municipalities also hosted significant non-irrigated agricultural activity, inadvertently leading to the inclusion of non-irrigated areas being classified as irrigated in the resulting map. The mapping accuracy was compromised due to the overlap with non-irrigated regions.

With the adoption of the new approach in Collection 9, a substantial enhancement in mapping accuracy has been achieved. This enhancement is attributed to the utilization of a regular grid provided by the ANA covering the semi-arid regions with irrigated agriculture. Each grid cell measures approximately 0.20° x 0.20° in size. This new grid-based approach has helped to avoid the previous issue of misclassification of non-irrigated areas. Figure 10A presents a comparison between the region adopted in the previous Collections and the new grid-based region adopted in Collection 9 to map the 'Other Irrigation System' class. This transition to the grid-based methodology has resulted in improved accuracy and a more precise representation of irrigated areas.

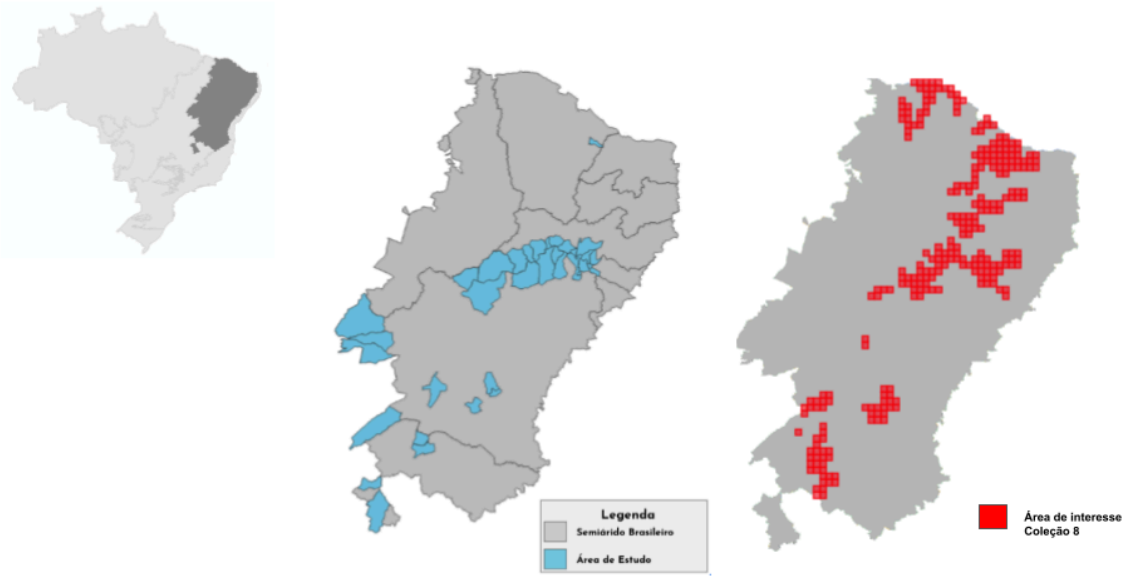


Figure 10A Comparison between the region adopted in the previous Collections and a new region adopted since Collection 8 to map ‘Other Irrigation System’ class.

4.3 Classification

4.3.1 Classification scheme

The classification process for the 'Other irrigation systems' class involves the consideration of two main groups: 'Irrigated agriculture' and 'Non-irrigated agriculture'. To enable the mapping of this class, training samples are obtained from the reference map provided by the ANA. These training samples cover both 'Irrigated agriculture' and 'Non-irrigated agriculture' regions and are used as training for the Random Forest classifier. This trained classifier is

subsequently used to identify areas of irrigated agriculture within the annual Landsat mosaics. This classification procedure is carried out exclusively within the geographical boundaries set by the ANA's GRID, ensuring accuracy in identifying irrigated agricultural regions.

4.3.2 Feature space

For the mapping of 'Other irrigation systems' alongside the data obtained from Landsat program satellites, supplementary metrics and indices were also calculated to enhance the identification of irrigated agricultural areas.

Table 3A presents the set of annual metrics used to map irrigated agriculture in the Brazilian semi-arid region.

Table 3A. Set of metrics used to map irrigated agriculture in the Brazilian semi-arid region.

Source	Bands and Spectral indices	Metrics
Landsat	BLUE	EVI2 Quality Mosaic
	GREEN	Minimum
	RED	Maximum
	NIR	Median
	SWIR1	Standard Deviation
	SWIR2	
	TIR1	
	EVI2 (JIANG et al, 2008)	
	NDWI (GAO, 1996)	
	MNDWI (XU, 2006)	
	CAI (NAGLER et al, 2003)	

4.4 Classification algorithm, training samples and parameters

The methodology employed for mapping irrigated agriculture within the semiarid region was founded on the application of the Random Forest algorithm. It used annual Landsat image

mosaics, incorporating Landsat spectral bands—specifically, BLUE, GREEN, RED, NIR, SWIR1, SWIR2, and TIR1. These spectral bands yielded valuable insights into both physical and biological surface attributes, facilitating the accurate differentiation of distinct land cover classes. Furthermore, statistical metrics such as Minimum, Maximum, Median, and Standard Deviation were computed for each spectral band, along with the EVI2 Quality Mosaic. This augmentation aimed to refine the spectral signal of the target class. Additionally, a selection of vegetation indices, EVI2, NDWI, CAI, and MNDWI, were integrated into the analysis to further enhance the classification process.

To initiate the process, 10,000 training samples were collected for both the 'Irrigated agriculture' and 'Non-irrigated agriculture' classes, considering the reference map provided by the ANA for the year 2019. These training samples were strategically acquired within the designated grids of the irrigated agriculture region in the semiarid region, as outlined by the ANA.

The Random Forest model was trained using these training samples and the Landsat mosaics, with a total of 100 trees in the model. The classification procedure was exclusively confined to grids demarcated by the ANA, aligning with their predefined geographical boundaries. The assimilated training samples corresponded to the ANA's reference map, guaranteeing the accurate representation of 'Irrigated agriculture' and 'Non-irrigated agriculture' across the region.

4.5 Post-Classification

The post-classification process of irrigation agriculture maps included the application of temporal and spatial filters.

4.5.1 Temporal filter

In the other irrigation systems mapping, a moving five-year window was also used, but using a different rule from the center pivot irrigation systems. In this filter, if the evaluated pixel was in the same class as at least three other pixels (previous, ahead or both), it remains in that class. However, if the evaluated pixel was not of the same class as at least three pixels (previous, ahead or both), the class was changed.

4.5.2 Spatial filter

In the other irrigation systems, a spatial filter was used to remove pixels that had less than 6 other connected pixels.

4.6 Results

When evaluating the 'Other irrigation systems' class mapped by MapBiomas Collection 11 and comparing it with ANA and IBGE data, discernible trends and disparities become evident (Figure 11A). This examination underscores the variability in reported values among the datasets, which may arise from differing data sources, methodologies, and assessment scopes.

According to MapBiomas Collection 11, the temporal evolution of the 'Other irrigation systems' class reveals fluctuating patterns across the years. The mapped area starts at approximately 0.015 Mha in 1985, decreases during the following years, and subsequently shows a gradual increase, reaching approximately 0.047 Mha in 2009. Between 2010 and 2018, the mapped area remains relatively stable, ranging mainly from 0.039 to 0.047 Mha. A more pronounced increase occurs from 2019 onward, when the mapped area rises from approximately 0.107 Mha to a peak of 0.176 Mha in 2020. After this peak, the area shows a slight and gradual reduction, reaching approximately 0.168 Mha in 2025.

In contrast, ANA data provides information for specific years, indicating an area of approximately 0.344 Mha in 2015 and 0.318 Mha in 2019. Similarly, IBGE data, available for 2017, records an area of approximately 0.262 Mha for the 'Other irrigation systems' class.

While the datasets exhibit an overall upward trend, particularly in the MapBiomas temporal series, the variations between reported values reflect differences in data collection, processing, and methodologies. It is noteworthy that ANA and IBGE data are collected at specific intervals, whereas MapBiomas Collection 11 provides a continuous temporal perspective. Moreover, even after the marked increase observed in MapBiomas from 2019 onward, its mapped areas remain lower than the values reported by ANA and IBGE. These disparities emphasize the importance of evaluation and cautious interpretation when utilizing such datasets to comprehend land-use dynamics and trends in the 'Other irrigation systems' class.

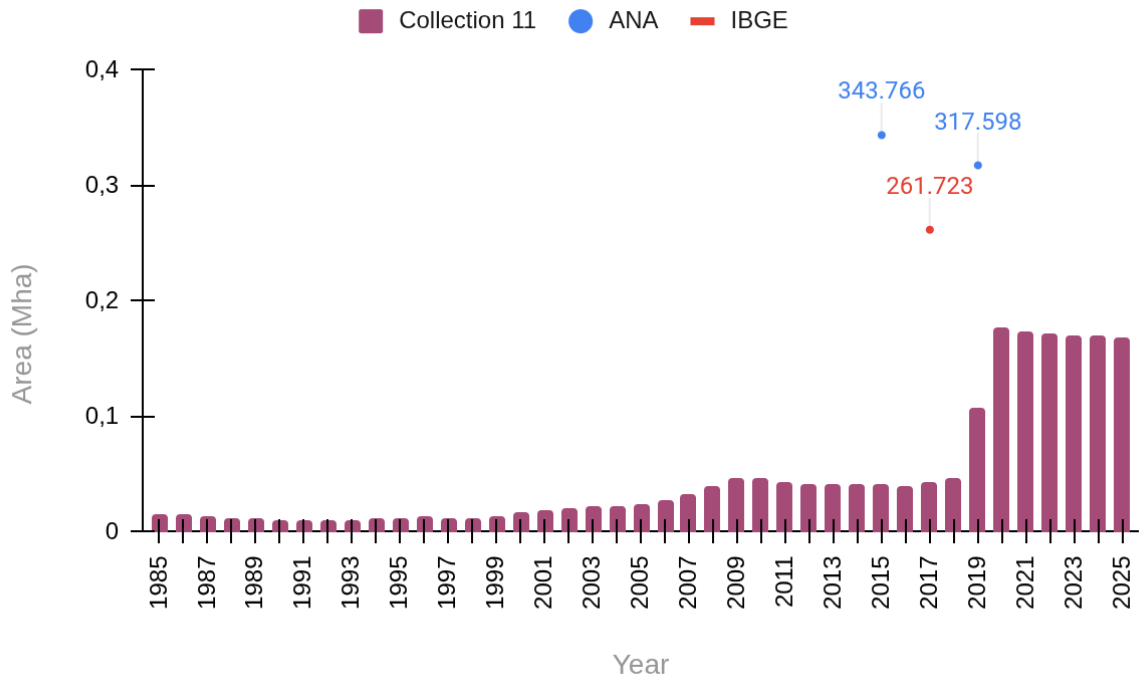


Figure 11A. Results of the 'Other irrigation systems' class in the semiarid region for the period from 1985 to 2024 compared to surveys carried out by the *Atlas da Irrigação* (ANA, 2017, 2021a) and the *Censo Agropecuário* (IBGE, 2009, 2019).

5 Center pivot irrigation information (*Discontinued*)

This product is discontinued and was last updated in Collection 9. The individualization of center pivot irrigation was a first attempt at understanding the dynamics of brazilian agriculture, using an object-based approach. Many of the advancements made developing this methodology were used to develop the Number of Cycles product, which aims to identify similar information in a pixel-based approach.

Understanding center pivots irrigation dynamics allows us to improve our understanding about most parts of irrigated agriculture in Brazil. The first effort to understand irrigation systems in the MapBiomas project began in Collection 5, with the use of innovative methods of Artificial Intelligence, through convolutional artificial neural networks to perform semantic segmentation of pivots throughout the Brazilian territory. In Collection 6 there was an expansion of the years mapped, with generation of time span maps of the entire MapBiomas series. In Collection 7.1 in addition to another time series expansion of center pivot irrigation map, covering from 1985 until 2021, it was also made effort to improve our understanding about crop dynamics in center pivot irrigation. Then, in Collection 7.1, a methodology was developed to provide more detailed information about this system, such

as the number of cycles performed per pivot in the crop year, the dates of start and end of each cycle, in addition to information about accumulated precipitation in each pivot and each crop cycle, initially only for Minas Gerais state between 2015 and 2021.In Collection 8 and 9, the pivot dynamic was extent to all of Brazil, and the 2022 and 2023 years, respectively, was also processed, providing dynamic information about the Brazilian pivots from 2015 to 2023.

5.1 Overview of the classification method

To provide this information to each pivot, several steps are necessary, from applying a Deep Learning model for center pivots irrigation individualization to obtaining smoothed time curves to identify the number of annual cycles existing in each of these pivots. Figure 12A presents all steps of this methodology.

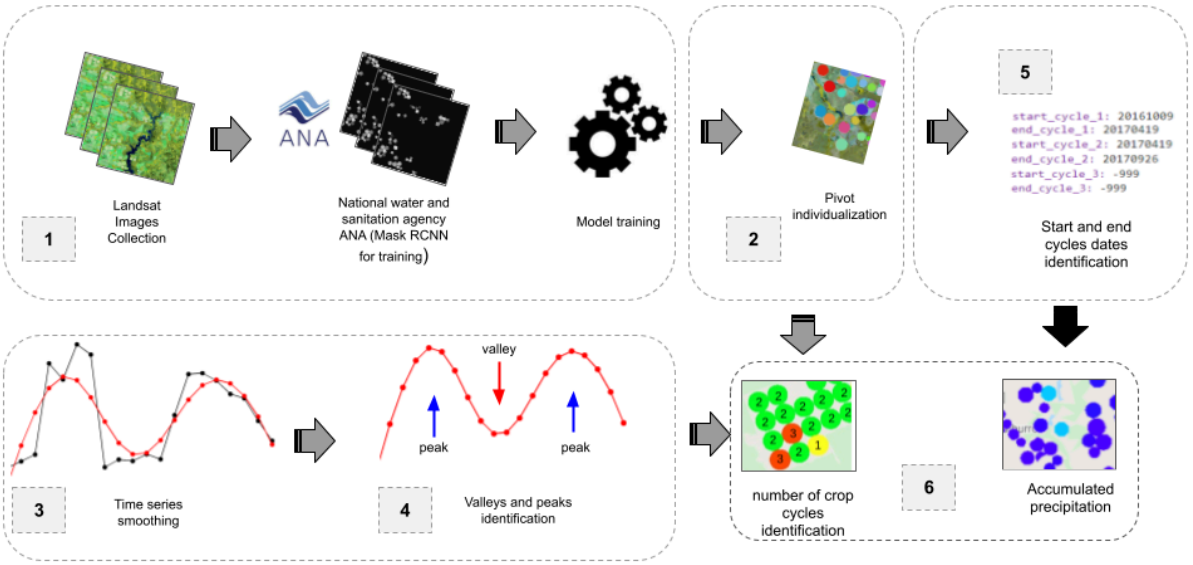


Figure 12A. Flowchart of the methodology necessary to obtain the number of cycles per pivot. 1) Obtaining the training dataset for the neural network (Landsat image and Mask of pivots); 2) Training the Deep Learning model for individualization of the pivots; 3) Landsat time series curve smoothing; 4) Identification of peaks and valleys in the curve; 5) Identification of the start and end dates of each cycle; and 6) Obtaining number of crop cycles and accumulated precipitation per cycle per pivot.

5.2 Center pivot individualization

5.2.1 Image selection

To individualize each center pivot irrigation, we used annual mosaics generated from available images for each year. Therefore, images from the Landsat series were obtained on the Google Earth Engine platform (Collection 2 Tier 1 TOA) in the period of 2015 to 2023. Only images with under 80% cloud cover and shadows were considered.

5.2.2 Definition of regions for classification

To individualize each center pivot irrigation, samples were first selected that represent relevant information about the pivots, so it was decided to select the blocks that contained at least 5 pivots. Thus, the samples were stratified between test areas (blocks with at least 5 and at most 9 pivots) and training areas (blocks with more than 10 pivots).

Figure 13A presents the blocks used for the RCNN (Region Based Convolutional Neural Networks) Mask prediction for all Brazil. The Mask R-CNN is a Convolutional Neural Network (CNN) and state-of-the-art in terms of image segmentation. This variant of a Deep Neural Network detects objects in an image and generates a high-quality segmentation mask for each instance.

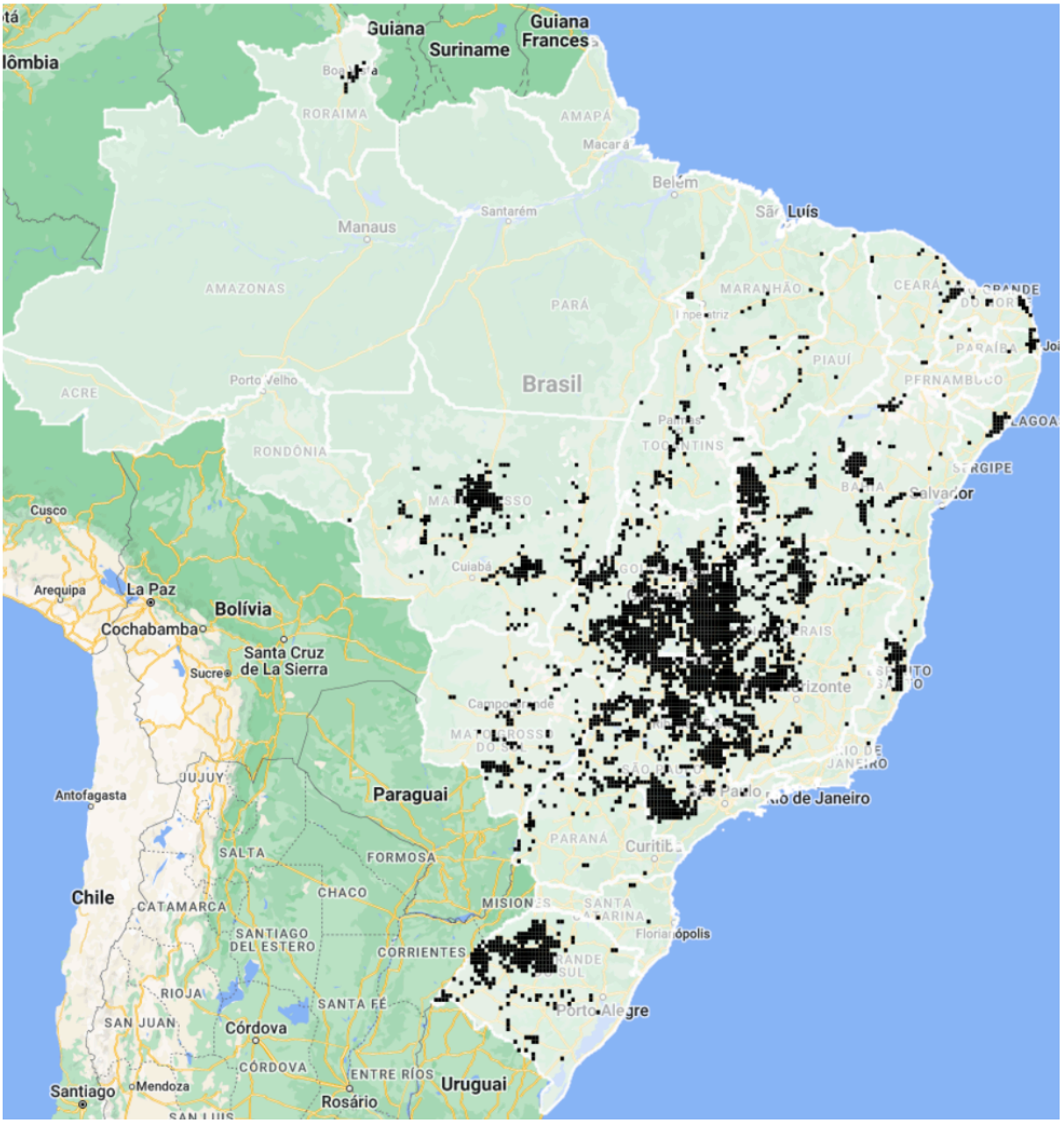


Figure 13A. A) Tiles (grid in black) for RCNN Mask prediction. Note: Areas without tiles indicate the non-existence of irrigation center pivots, according to the map made available by ANA for the year 2019.

5.2.3 Classification

5.2.3.1 Classification scheme

The RCNN was trained to identify and individualize each center pivot irrigation. Semantic segmentation considers two classes, (binary classification), 1 for ‘pivot’ and 0 for ’non pivot’. In instance segmentation, however, each pivot is mapped separately, adding one unique ID for each.

5.2.3.2 Feature space

The Normalized Difference Vegetation Index (NDVI) (ROUSE et al., 1974) was calculated for each image in order to generate standard deviation and percentiles metrics, as presented by Table 4A. These metrics were chosen seeking to capture not only the temporal variation of NDVI in the pivots, but also the variations of other agricultural targets outside of pivots (such as pasture, barren soil, native vegetation, etc).

Table 4A. Index and metrics used to individualize center pivot irrigation.

Indexes	NDVI
Metrics	stdDev, 75th percentile, 100th percentile

5.2.3.3 Classification algorithm, training samples and parameters

Instance segmentation is performed from a pre-trained neural network of Mask RCNN type architecture. This architecture was developed in Python, using the Pytorch framework, along with the Detectron 2 package. Figure 14A represents the flowchart of the entire Mask RCNN training process.

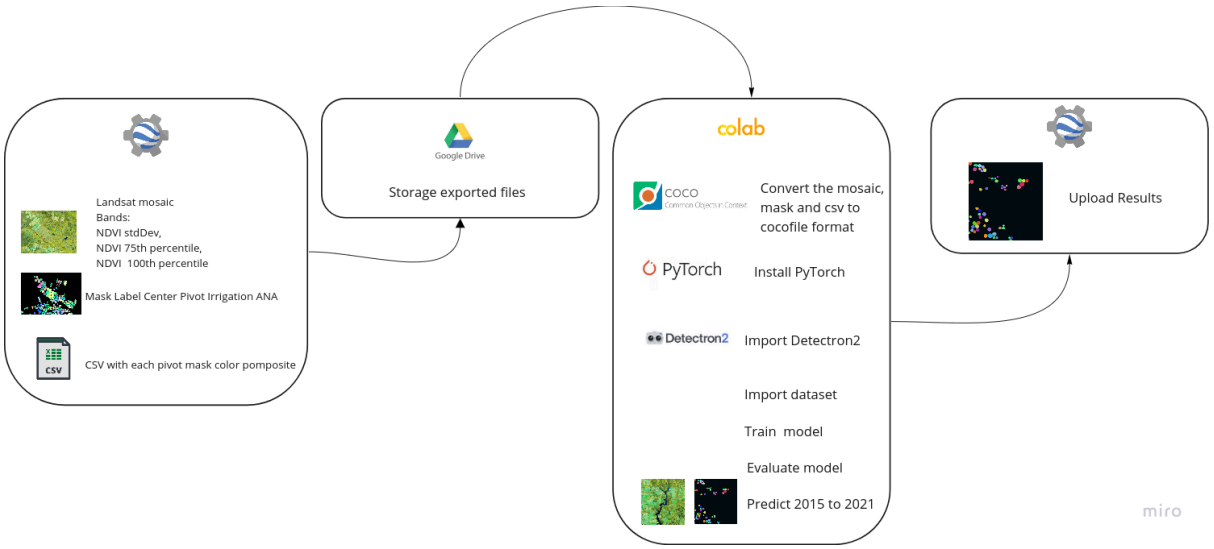


Figure 14A. Pipeline to use Detectron2 to pivot instance segmentation.

5.2.4 Post-Classification

Post processing of the center pivot irrigation has two more steps besides the spatial and temporal filters, which are focused for solving pivots 'union' and 'edge' problems.

5.2.4.1 Union Problem

Union problem consists of a false pivot generated between real pivots that are overlapping. Figure 15A exemplifies this problem as well as its resolution.

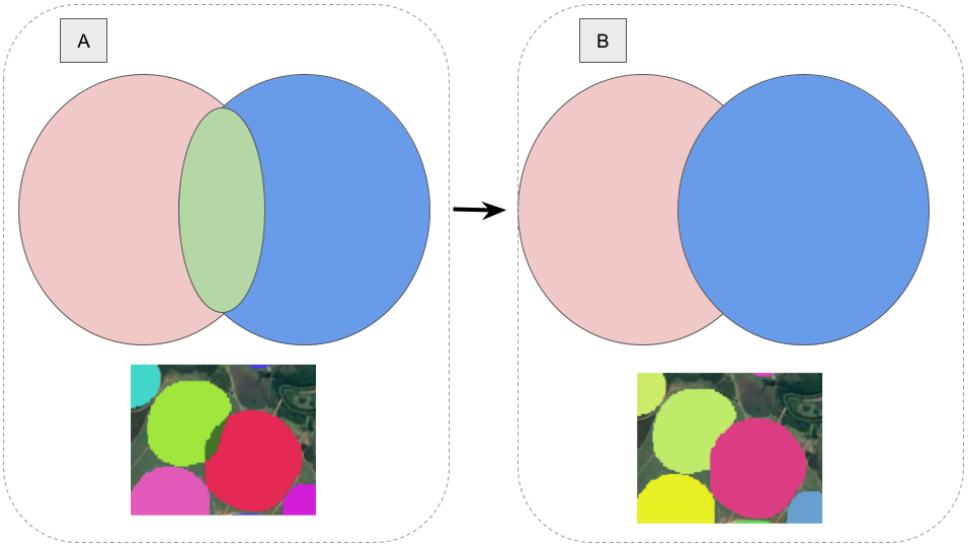


Figure 15A. Illustration of the problem of the union between two or more center irrigation poles. A) union problem; B) result of the filter applied to solve the union problem.

To solve this union problem it is necessary to find the ID of the false pivot (generated by union of two or more pivots through a sum of true IDs), and then identify the IDs pivots that generated this false pivot ID. Based on this information it is possible to replace a false ID to a true ID, from one of pivots that generated this false ID.

5.2.4.2 Edge Problem

The edge problem is a result of the shape and size of the RCNN Mask input. Some pivots will inevitably be "cut off" due to the size of the input tiles, i.e. one part of the pivot will be in one block (tile) and the other part will be in another adjacent block. The edge problem was solved with the application of two complementary filters (erosion and dilation) and a reduction by spatial connectivity. Figure 16A shows an example of an application to solve edge problems (A).

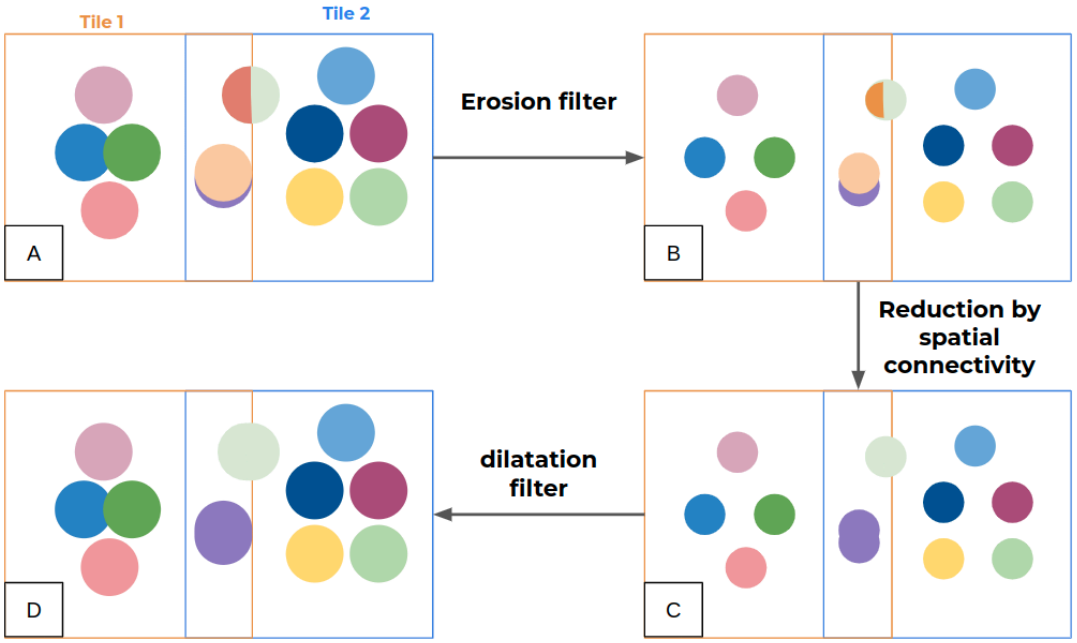


Figure 16A. A) Example of application of morphological filter to solve edge problems. A) tiles with pivot edge problems; B) erosion filter; C) spatial connectivity Reduction and D) dilatation filter.

The erosion filter is applied to the tiles in order to reduce the size of the instances (pivots) with the goal of isolating them from each other (B). Then, pivots that were "cut off" by the tile of the RCNN Mask and that exist in the overlap of both boundary images are connected (C). Finally, after this step it is necessary to apply the morphological dilation filter to return pivots to their original size (D) .

5.2.4.3 Temporal filter

For temporal consistency of the IDs over time, a temporal filter was applied with the goal that each pivot remains with the same ID over the years. In this step, a reference image was generated through the accumulation function of all years (2015 to 2022), thus the reference image has all the pivots of the time series and their respective IDs. An accumulation of pivots must be calculated for each year, for example, the accumulation of the year 2020 has the pivots of the years 2015, 2016, 2017, 2018, 2019, and 2020.

5.2.4.4 Spatial filter

A spatial filter was used to remove areas smaller than 10 hectares, so that the noise caused by the accumulation function is excluded, as shown in Figure 17A.

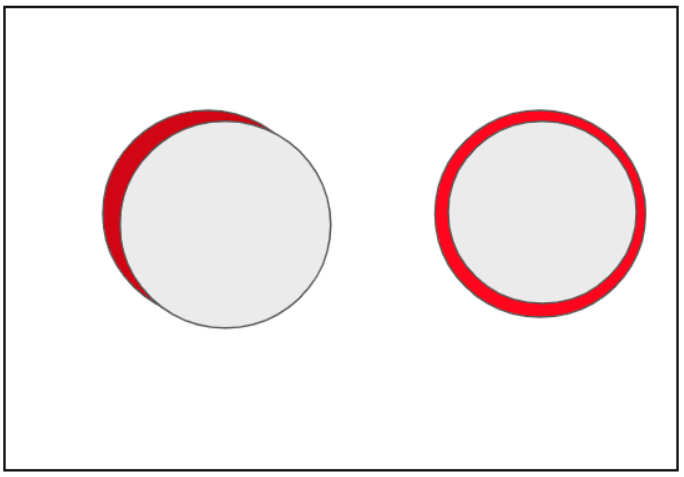
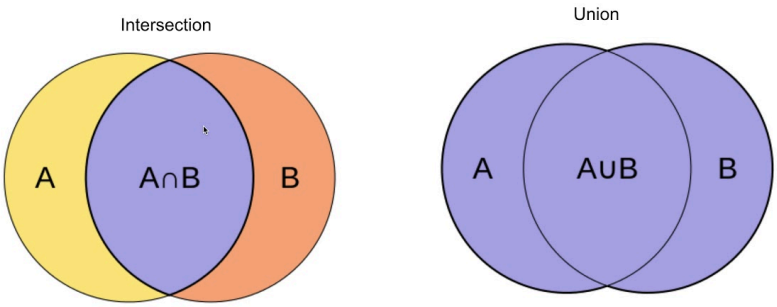


Figure 17A. Example of spatial filtering. Pivot polygons in red noise generated by the accumulation function that are removed through the spatial filter.

5.2.5 Validation strategies

To validate the instance segmentation (Mask RCNN) and semantic segmentation (Unet) methodologies the Jaccard index was calculated. The Jaccard index (JACCARD, 1901), also known as the Jaccard similarity coefficient or *intersection over union* (IOU), is a statistic used for gauging the similarity and diversity of sample sets and is defined as the size of the intersection divided by the size of the union of the sample sets (Figure 18A).



$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

Figure 18A: Index Jaccard (IOU) calculation.

The spatial similarity of the Unet results with the Mask RCNN results obtained for the year 2019 was 61.7%. The location data of the center pivots of public irrigation by ANA showed 63.1% similarity with the results obtained from the RCNN Mask. The similarity between ANA and Unet data was 78.4%. It is important to note that instance segmentation is a new methodology that is still under development.

5.2.6 Results

Mask RCNN returns as output a raster of the input mosaic size (15 x 15 km) composed of 0, which corresponds to no detection of center irrigation pivots and values corresponding to the ID of the classified pivots. Figure 19A shows the input and output of the RCNN Mask prediction.

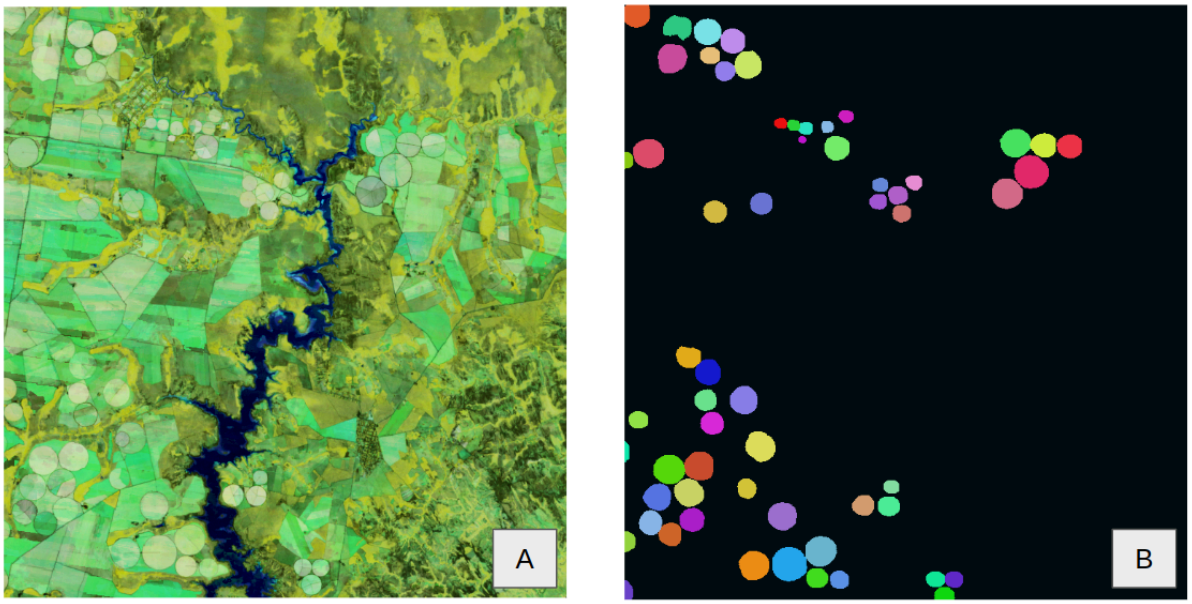


Figure 19A. Result of the Mask RCNN prediction.

5.3 Center Pivot Information

Crop and environmental characteristics of center pivot irrigation were obtained to each individual pivot. Thus, using the geometry of each pivot, the following information was extracted: *i*) crop cycles number of each pivot; *ii*) start and end dates of each cycle; *iii*) crop cycle length in days of each cycle; and *iv*) daily average precipitation of each cycle. Besides this, it was also possible to obtain the information if the pivot was in a perennial cultivation area, and if it did not present internal coherence – due to multiple crops at the same time or complex management – , it was not possible to obtain the previous information (*i*, *ii*, *iii* and *iv*) and the pivot is defined as a non-classified.

5.3.1 Image selection

The period used to select the images to obtain a temporal EVI2 curve was based on the crop year. The crop year is different from the conventional year (from January to December), since the crop year aims to define the period when the cultivation occurs in a determined region. Thus, depending on the type of agriculture, the crop year can start in any month of the year, generally following the rainy season, since in this period there is humidity available to the crop development. Thus, to attribute information for each center pivot irrigation,

Landsat images (TOA) were selected from a crop year, in order to compute the EVI2 time series.

The crop year was defined automatically for each Landsat scene and year. We defined the crop year as 3 months before and 9 months after the mean vegetative peak month of each scene, based on an EVI2 curve of MODIS observations.

5.3.2 Method to attribute information to each pivot

5.3.2.1 Crop cycles number

The first information obtained, that is a base to obtaining the others, was the number of crop cycles of each pivot. An EVI2 curve of Landsat images from the crop year was smoothed to minimize noise and to reconstruct the time series. The Whittaker method (WHITTAKER, 1922) was used to smooth EVI2 temporal series, since this method presents a great alternative to smooth and to reconstruct temporal series, most importantly keeping only meaningful variations and preserving the temporality of them.

Based on the smoothed EVI2 curve, it was possible to identify when inflections occur in the curve, that is, the change of direction of the curve. Thus, it was possible to identify the points of valleys (defined as the inflections of change from negative to positive sign), and peaks (defined as the inflections of change from positive to negative sign) (Figure 20A). Finally, to define the number of cycles, this can be counted as the number of peaks (or valleys minus one), determining the number of crop cycles in a period.

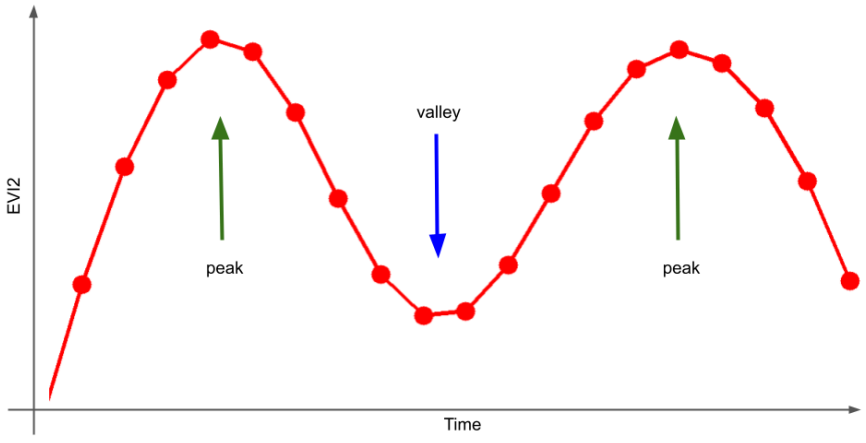


Figure 20A: Peaks and valleys identification based on the smoothed EVI2 time series curve.

5.3.2.2 Start and end cycles dates

After identifying the peaks and valleys over the EVI2 time series, it was possible to determine the start and end dates of each cycle. In this step we sought to identify the dates of the valley inflections of the EVI2 curve of each pivot. According to the amount of Landsat

images available in the crop year, where each one represents a binary information of valley (1) or no-valley (0), it obtained a percentage of presence of pixels identified as valley (1). Then, after valleys date identification it was checked if the quantity of them was equivalent to the expected quantity for the number of cycles of the pivot (number of cycles +1). If this information is true, the pivot is considered well identified, if not, the pivot is considered as non-classified, since due to the internal dynamics (this usually occurs for pivots with different crops at the same time) it was not possible to identify a spatial coherence in the start and end date definition.

The valley dates were used as a base to determine start and end cycle dates. Based on the daily interval between the two valleys that compose a cycle, the start date was defined as the 20th percentile value, and the end date as the 80th. This was done to reduce cycle coverage to the period where the crops were active, eliminating soil management periods (JÖNSSON and EKLUNDH, 2004). To avoid omitting the planting period, a -15 days buffer was also added to the start cycle dates.

This cycle delimitation method has some known issues, such as the delimitation of cycles based only on the time value. Improvements in this area will be sought for future collections.

5.3.2.3 Crop cycle length

Crop cycle length in days was obtained as a difference between the end and start date of each pivot for each cycle.

5.3.2.4 Average Daily Precipitation

For Precipitation information we used data from Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) product (FUNK et al., 2015), that provide daily and sub-daily precipitation information for quasi–global spatial coverage (50°S-50°N), from 1981-present, in a 0.05° x 0.05° of spatial resolution. Based on CHIRPS data it was obtained an accumulation precipitation of each pivot and then this amount of precipitation was divided by the number of the days of each cycle, resulting in an average daily precipitation per cycle.

5.3.2.5 Additional Information

In addition to the number of crop cycles, start and end dates, and precipitation information, the product also provides additional information about non-classified, perennial and sugarcane pivots.

Perennial and sugarcane pivots were identified using the respective maps from MapBiomas Collection 9, since cycles and environmental information were only accounted for temporary crop pivots.

Pivots in which it was not possible to identify the start and end dates of each cycle were set as non-classified. This problem can be due to a number of factors. For instance, pivot internal crop dynamics, when there are multiple crops on a single pivot, or the same type of crop, however at different times. There is also the possibility of errors inherited by the individualization of pivots methodology, since a not well-defined geometry may encompass other land uses or surrounding crops. In these situations, it was not possible to identify coherent start and end dates, since there is no agreement inside the pivot geometry. Figure 21A presents some examples of when it is possible to identify start and end dates and when this identification is not possible, resulting in non-classified pivots.

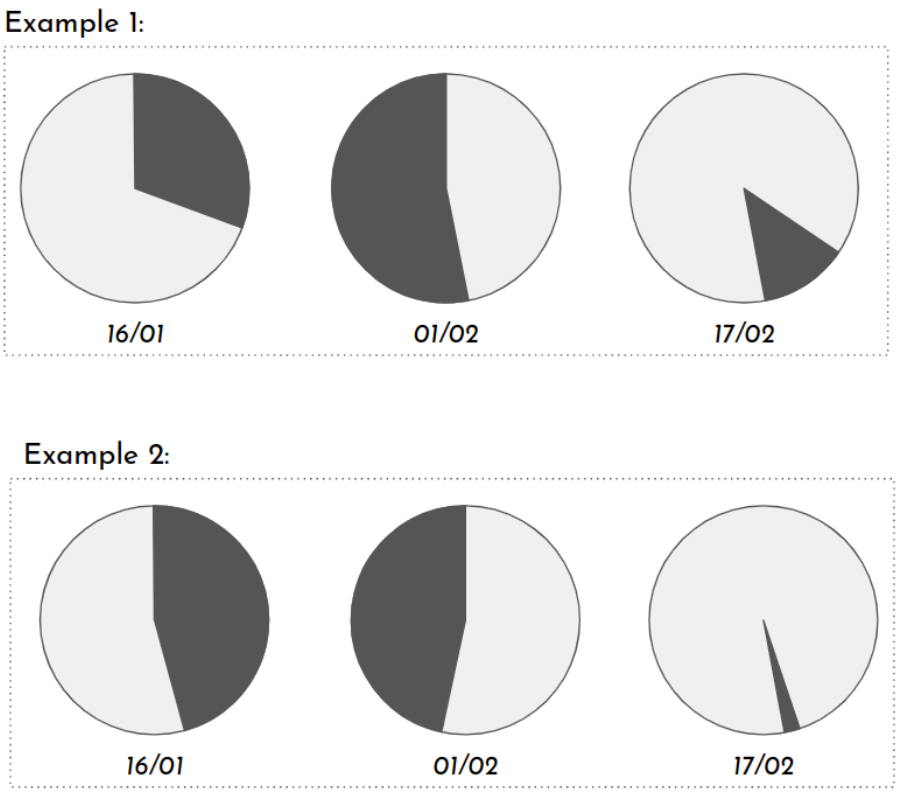


Figure 21A: Examples of start and end dates identification, and limitations that cause non-classified pivots.

Example 1 shows a pivot in three different Landsat dates. The shaded area represents the area identified as a valley date. In the first image (16/01) there are approximately one quarter of the pivot identified as a valley. This amount is even less in 17/02. However, in the second date (01/02), most of the pivot was identified as a valley, providing a valley date information.

Example 2, on the other hand, presents a more complex situation. In these three dates (16/01, 01/02, and 17/02), there is not a single image where most of the pivot is identified

as valley. In this situation there is no possibility to obtain start and end dates by the same method as before (statistical mode), so pivots in this or similar situations were set as non-classified. This is a known flaw in the methodology and improvements will be sought in future collections.

5.3.3 Validation strategies

From the total of classified pivots, 250 of them were randomly selected and evaluated year by year, visually, in order to validate the consistency in terms of class, start and end dates of each cycle, and also the consistency of number of cycles. This total (250) represents between 3 to 6% of total pivot amount, depending on the year, since the number of pivots increases over the years.

The following errors were considered: (1) cycle number errors, when a pivot shows a difference between the number of cycles identified in the methodology and the visual analysis; (2) dates errors, when the pivot has at least one cycle where the start and end dates are not consistent with the expected in the vegetation index curve; (3) class errors, where the pivot was misclassified in any way. The first two errors can occur at the same time, and when so it is likely that the pivot was mostly not well defined. However, the errors individually do not indicate that the pivot is entirely wrong. Class errors can be associated with an omission from the sugarcane and perennial masks. Cycle number and date errors individually show that a cycle in the pivot was misidentified in some way, but not necessarily all of its cycles. Figure 22A presents the rates (%) of pivot without any kind of error, with only one error and with two errors.

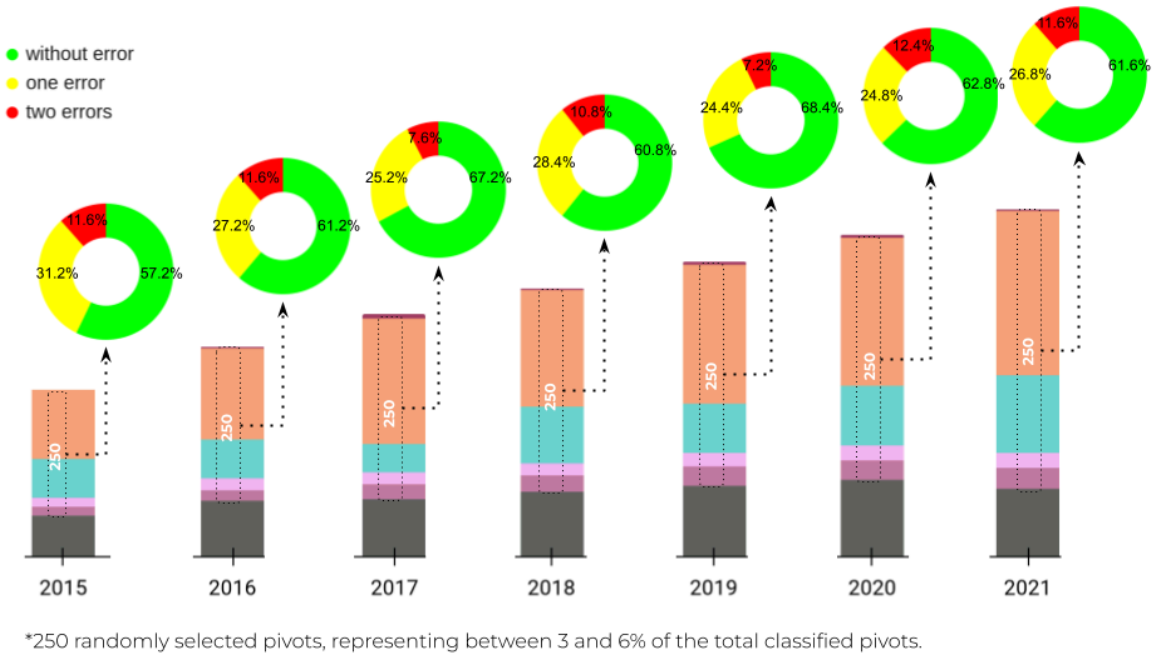


Figure 22A: Percentage of pivots without errors, with only one error and with two errors identified in 250 pivots selected randomly.

The results presented in the Figure 22A above, show that about 57.2 to 68.4% of pivots randomly selected have none of the analyzed errors, while around 24.4 to 31.2% of samples presented only one error, and 7.2 to 12.4% of data evaluated presented two types of errors. The analysis also provides information about the type of these errors identified. Figure 23A presents the error rates identified by the pivots sample randomly selected.

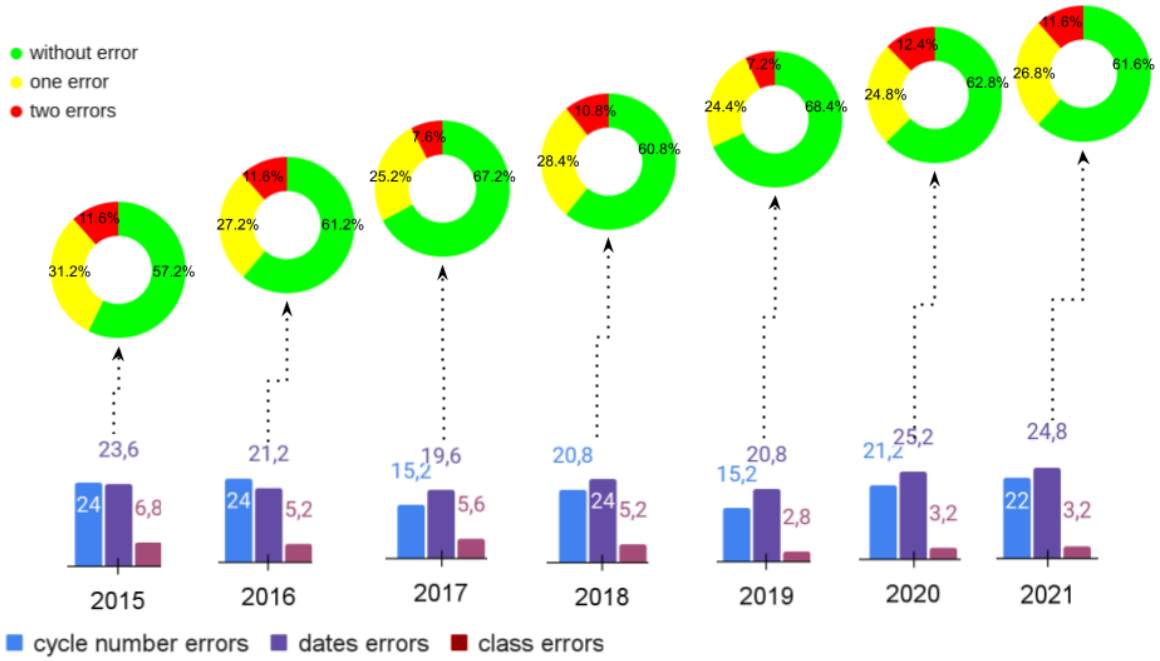


Figure 23A: Percentage of type of errors identified in 250 pivots selected randomly.

Figure 24A informs us that we have two main types of errors on pivots. For instance, errors due to incorrect cycle accounting are about 15 to 25% of the errors identified. About 19.6 to 25.2% of errors are related to errors in identifying the start and end dates of cycles, and less than 7% of errors are related to errors in pivot class, i.e. classification of the type of use and coverage of that pivot.

5.3.4 Results

The results presented in Figure 24A highlight about the example about the expansion of the pivots over Minas Gerais, detailing about the intensification of the agriculture through the irrigation, showing the number of pivots, for each year, with 1, 2 and 3 cycles.

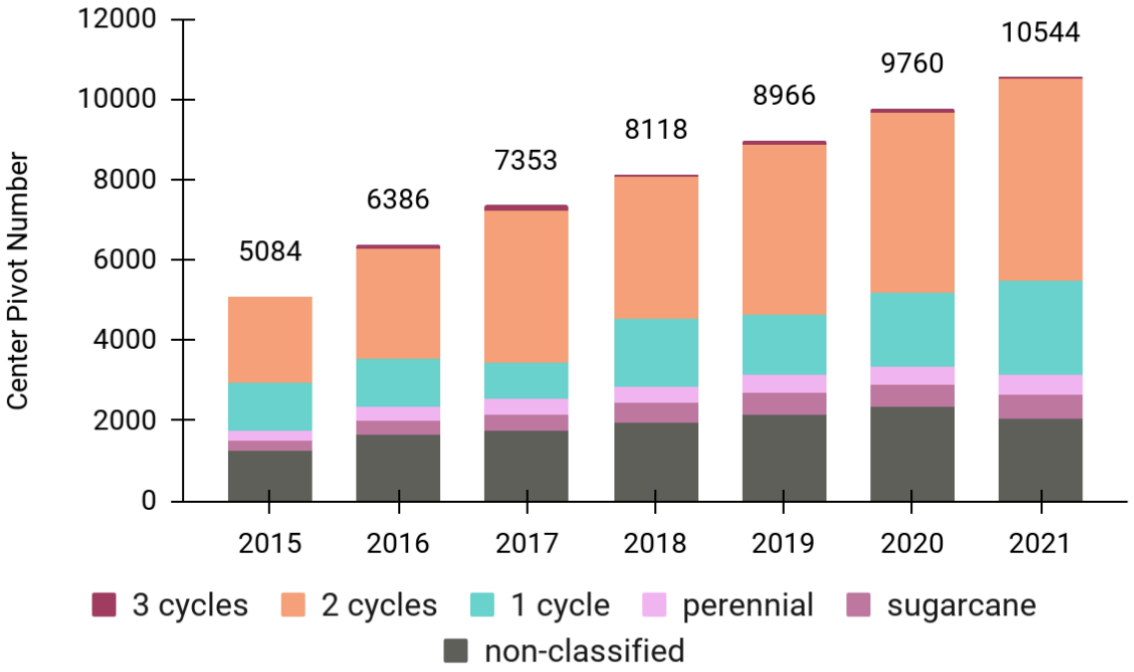


Figure 24A: Number of center pivot irrigation with one, two and three cycles, perennial, sugarcane and non-classified in Minas Gerais state.

Overall, in Brazil, 19% of center pivot irrigation has one crop cycle, while 46% present two crop cycles, and approximately 1% were cultivated three times over the analyzed years (2015-2022). In addition, approximately 24% of center pivot irrigation was not classified as a function of methodology limitation. Around 6% of these pivots were identified as perennial pivots and 5% as sugarcane.

In addition to this information on the expansion and intensification of agriculture, in terms of the number of gullies, some other information on the dynamics of gullies can also be obtained from this product. For example, for the state of Minas Gerais, some general information about the region's crop dynamic was summarized. It was possible to identify the months of start and end cycle, the number of days of each cycle and how much precipitation occurred in this period, also to each cycle. Figure 25A presents a temporal average of this information for the pivots identified with crop dynamics (pivots with 1, 2 or 3 cycles), in the Minas Gerais state.

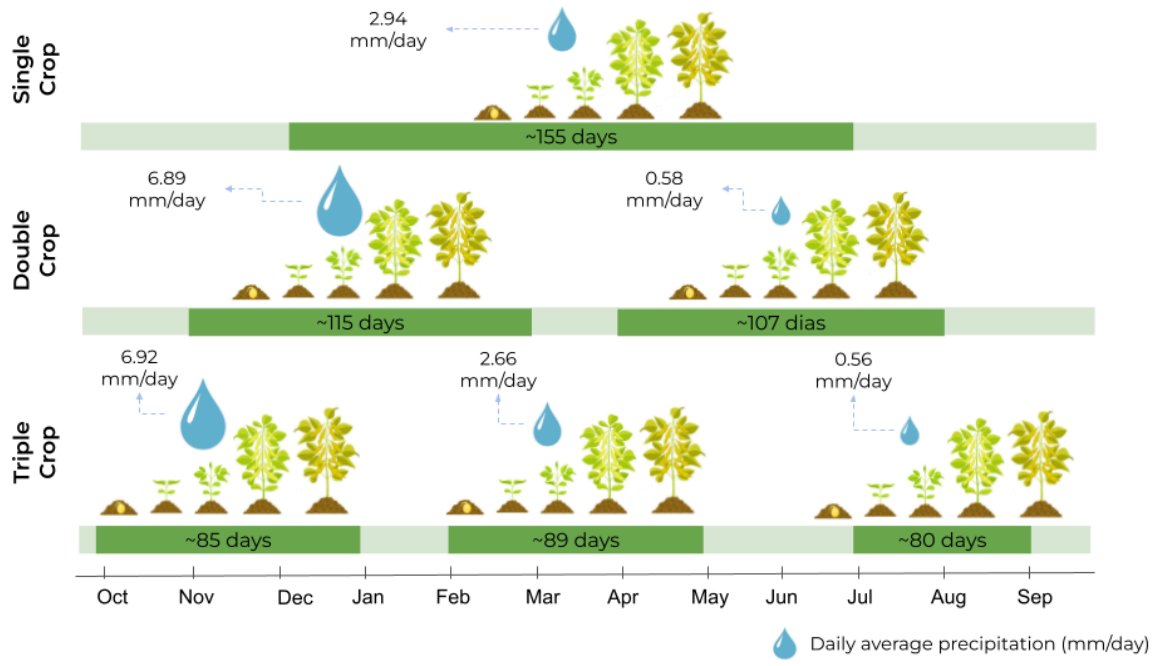


Figure 25A: Summary of pivot dynamics for Minas Gerais state.

In Figure 25A presented above, it is possible to see that both the start and end months of each cycle, as well as its duration and daily average precipitation varies according to the number of times the pivots have been cultivated in the year.

For instance, on pivots with one single crop cycle, the cultivation usually starts in December with a harvest occurring around July. Pivots with only one cycle present a longer cycle duration, of approximately 155 days,. The daily precipitation is about 2.94 mm/day, since the cycle also extends over rainy (December, November, January and February, June and July) and dry (May, June and July) months in this region.

For pivots that present a double-crop there is a different dynamic. In these pivots the cycle duration is around 115 days for the first cycle and around 107 days for the second cycle. Comparing pivots with single and double-crops it is possible to verify that for double-crop pivots the period when the cultivation occurred was usually earlier than the single-crop pivots, with planting starting around November and harvesting around March. As this first cycle of the double crop pivots is concentrated in the rainiest months of this region, the daily average precipitation is higher, around 6.89 mm/day. Regarding the second cycle of double-crop pivots, this usually starts planting around April and harvesting usually occurs in August. Because it comprises mainly the winter months, this period presents low precipitation, with a rate of approximately 0.58 mm/day, indicating that only the precipitation of the period is not sufficient for the development of the crop, requiring the use of an irrigation system for this second harvest.

For pivots with triple-crop the crop duration is even shorter than pivots with one or double-crops. In these pivots with triple-crops, generally the cycles extend from from 80 to

89 days. In addition, the period of cultivation of the first cycle starts earlier than pivots with single or double-crops, starting in October and the harvest occurring in January. The second cycle normally starts in February and the harvest occurs in May, while the third cycle extends from July to September. The first two cycles of triple-crop pivots take place at least partially in the region's rainy season, with a rate about 6.92 mm/day for the first cycle and 2.66 mm/day for the second cycle, indicating a lower dependency of the irrigation system in the first cycle compared to the second one. However, during the third cycle precipitation rate is about 0.56 mm/day, suggesting a significant increase of irrigation importance for crop development.

B. Number of Cycles

1. Overview of the method

The number of cycles maps were a new product in MapBiomas Collection 9 that aimed to map the number of crop cycles in each temporary agricultural area in Brazil at the pixel level. These maps are produced based on an agricultural mask from the latest update of the MapBiomas 10m collection, which uses Sentinel-2 images at 10m spatial resolution. In this beta version, updated to Collection 10, annual maps are available from 2017 to 2024, always comprehending September of the previous year and ending in August of the target year. There is also a visualization by the mean number of cycles for all years, available in the Agriculture module inside MapBiomas' platform.

The method identifies the number of successive crop cycles in each pixel, from the start of detection of the plants' spectral response after emergence until the response declines with senescence, without distinguishing between crops. In this way, the product takes into account both the cycle of commercial crops for grain production and those intended only to produce biomass for ground cover in the off-season of the main crop. As a result, in some regions the area identified as having more than one cycle per year may be larger than that estimated by the official organizations that carry out crop surveys, as these only focus on the cycles of grain-producing crops.

2. Time-series construction

Sentinel-2 images were used to construct the time series, providing a temporal resolution of approximately five days. The images were acquired and processed using the Google Earth Engine (GEE). To highlight the spectral behavior of the vegetation, the Enhanced Vegetation Index 2 (EVI2) was calculated from bands 5 and 8, referring to the red and near infrared. The time series were constructed by crop year from 2017 to 2025, starting in September of the previous year and ending in August of the target year.

As it was desirable to have temporal consistency in the series and consequently not exclude any data, cloud and shadow masking was not applied. In order to remove noise and highlight cyclical behaviors, we opted to apply a smoothing technique to the series.

2.1 Time-series smoothing

The Harmonic Analysis of Time Series (HANTS) was used for smoothing, or more specifically an adaptation of the HANTS-GEE package. HANTS is a common technique for reconstructing time series and, like techniques based on the Fourier transform, it uses a decomposition of the time series into frequencies and considers high frequencies to be noise (Zhou et al., 2023). The parameters were defined based on tests with a visual assessment of the consistency of the results. Standard values were used, except for: number of harmonics of 5; adjustment error tolerance of 0.15; and maximum iterations of 2.

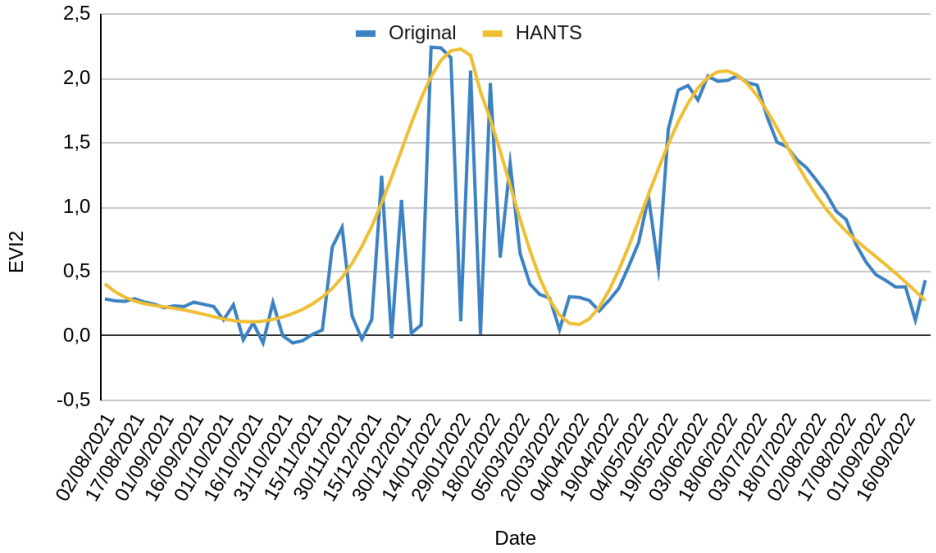


Figure 1B: Example of original EVI2 data and smoothing using HANTS in a series with two cycles.

3. Delimitation of crop cycles

To quantify the frequency of cultivation, an algorithm was developed to detect the occurrence of cultivation cycles based on inflections in the EVI2 curve. On a pixel scale, from the time series of values, the difference of each value in the time series with the immediately preceding and following value is calculated. Inflections in the curve are detected as a peak, when both differences are greater than zero, or as a valley, when both are less than zero. A crop cycle consists of a peak, which represents the greatest vegetative vigor of a period, and two valleys, which represent the beginning and end of the cycle, being an approximation of planting and harvesting.

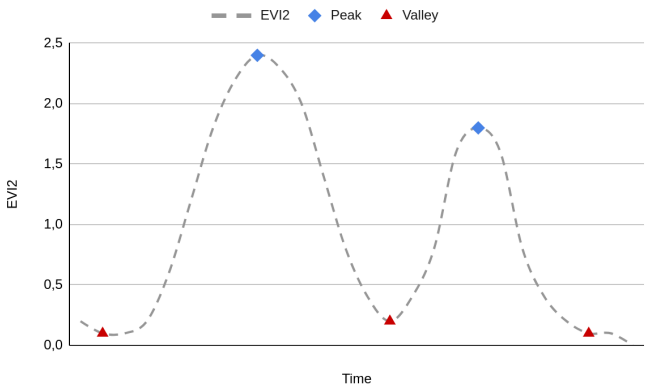


Figure 2B: Example of peak and valleys detection in a curve with two cycles.

Filters were applied to the detected inflections to avoid the influence of imperfections in the residual EVI2 curve from smoothing, which can generate false peaks and valleys. Two

parameters were provided to the algorithm: (1) minimum peak value, which is the lowest possible value for identifying an inflection as a peak, regardless of differentiation to adjacent values; and (2) minimum amplitude, or the minimum difference between a peak and the nearest valleys. These parameters were defined empirically as 1 for the minimum peak value and 0.7 for the minimum amplitude.

The frequency was calculated from the count of peaks detected in a series, which corresponds to the number of times that site was cultivated in the agricultural year, which usually varies from one to three times.

3.1 Spatial filtering

To reduce isolated noise and smooth the spatial distribution of the frequency classes at the pixel level, a neighborhood spatial filter was applied to the number of cycles map. The filter considers the immediately adjacent pixels using a Manhattan kernel with a radius of one pixel. The modal value within this neighborhood was calculated using the mode reducer, assigning to each pixel the most frequent frequency class observed locally. This procedure consolidates the predominant number of cycles in the surrounding area and reduces isolated pixel-level variations in the final map.

4. Accuracy evaluation

To check the accuracy of the mapping, the measured values were cross-referenced with the mapping of Center Pivot Irrigated Agriculture in Brazil (ANA, 2023), carried out by the National Water Agency (ANA). Among the various years mapped, the 2022 mapping identifies the dynamics of annual crops in all center pivot irrigation systems in the country, including, among others, pivots with one, two or three annual cycles. This data was chosen as a reference because it makes the same distinction as the result we intend to evaluate, although the scope of the validation is reduced to agriculture in this type of system only. The data is available free of charge from the ANA in vector format, and was rasterized with a spatial resolution of 10 m to make it compatible with the result to be evaluated. To avoid edge effects resulting from rasterization, pixels up to 50 m away from the edge of the central pivots were disregarded.

The evaluation was carried out in an area on the west region of Bahia state, corresponding to two Sentinel-2 imaging scenes: 23LLF and 23LMF. A stratified sampling of the reference map was carried out, where 500 samples were taken according to the proportion of each value. This totaled 182 points for one cycle, 270 points for two cycles and 48 points for three cycles. The points were cross-referenced with the results obtained for the 2021/2022 agricultural year, consistent with the period of the reference map, and a confusion matrix was constructed. Metrics were calculated for global accuracy (GA), producer accuracy (PA) and user accuracy (UA) per number of cycles.

	HANTS		
N. cycles	One	Two	Three
PA	0,87	0,99	0,85
UA	0,98	0,92	0,82
GA	0,94		

Table 1B: Accuracy metrics for the 2022 frequency map in the selected area.

5. Limitations

This beta version is a proof of concept of the method's viability on a large scale. Some inaccuracies are expected and can be a result of several known limitations, such as: crop-year being defined as the same period for the whole country, which in reality varies between regions; parameters being defined as the same for the whole country, which does not take into account possible variability of vegetative vigor (i.e. max and min EVI2) between biomes; no current way to discern if a detected cycle was a crop meant for commercialization or just a cover crop to manage the soil; lack of comparable data for a comprehensive accuracy evaluation. These problems are expected to be addressed in future revisions of the method.

C. Land Use Land Cover - Second Season

1. Overview of the method

This section describes the production of annual Second Season maps for corn in Brazil at 30 m spatial resolution from 2000 to 2025 using Landsat Collection 2 surface reflectance imagery. The workflow follows the main Land Use Land Cover (LULC) map's pixel-based paradigm (seasonal mosaics, engineered feature space, Random Forest classification, integration to LULC, and selective post-processing), but is tailored to second season phenology and implemented state-by-state. This section mirrors the Agriculture and Forest Plantation Appendix's organization and terminology while specifying the operational choices effectively adopted here.

We mapped second-season crops with emphasis on corn (primary target) and cotton, using a three-class training scheme: (1) corn, (2) cotton, and (3) aggregated other temporary crops. Classification was executed per state, not per Landsat scene, using a single statewide geometry in each case to standardize normalization and reduce edge artifacts across scenes. State-specific temporal windows (see below) were the only systematic variation across states.

Image selection windows were designed to capture the second-season vegetative peak. As a baseline, February–May was adopted in the Mato Grosso pilot and then adjusted by state according to the official agricultural calendar (CONAB) and visual inspection of monthly mosaics.

Input data comprise Landsat Collection 2 surface reflectance, processed over a single statewide geometry per Federative Unit (UF). Sensor allocation follows fixed rules: 2000–2002 rely on Landsat 7 (LE07) only; 2003–2011 on Landsat 5 (LT05) only; 2012 uses a merged LT05+LE07 stack; 2013–2025 employ Landsat 8 (LC08) every year, with Landsat (LC09) included where available (e.g., 2023; note 2013 = LC08 only, 2023 = LC08+LC09). For each year, scenes are constrained to the second-season window defined for the state (baseline 1 February–31 May), filtered by CLOUD_COVER_LAND < 40%, and masked using QA_PIXEL to remove cloud (bit 3), cloud shadow (bit 4), snow (bit 5), and dilated cloud/water (bit 2). Bands are harmonized to a common set—BLUE, GREEN, RED, NIR, SWIR1, SWIR2—by mapping bands SR_B1–B5,B7 for LT05/LE07 and SR_B2–B7 for LC08/LC09; EVI2 is computed per image and appended prior to seasonal compositing.

For each target year and state, we built a seasonal mosaic by reducing the filtered collection with median and percentiles (p20, p80) over the bands BLUE, GREEN, RED, NIR, SWIR1, SWIR2 and the EVI2 index. The resulting metrics were normalized with fixed clamps (same limits across years and states) prior to classification.

To stabilize the model across sensors/years, we adopted a two-year composite training for each epoch and applied a single Random Forest per epoch across the entire period:

- 2000–2012: RF(100 trees) trained with 2007 and 2012 samples, then applied to 2000–2012;
- 2013–2024: RF(100 trees) trained with 2013 and 2023 samples, then applied to 2013–2024.

Class labels were restricted to the three classes above; “other temporary crops” served to regularize decision boundaries between corn/cotton and non-targets within the second-season window.

For each year, the appropriate epoch classifier was applied to the normalized mosaic. Outputs were then restricted to agricultural areas using the MapBiomas Collection 10 integration map. The agricultural mask was implemented by remapping classes [39, 41, 62, 20] → [1, 1, 1, 0] (keep vs. exclude) before masking. This step follows the module’s integration logic for harmonizing class-specific products with the LULC map.

Spatial/temporal filtering was not universally applied. After targeted analyses, we executed a unified filter (eight-neighborhood connectivity; min-patch = 8 for 2000–2012 and = 20 for 2013–2024; class-wise temporal persistence rules) only in five states where it improved coherence. In other states, masked raw predictions were retained to preserve genuine interannual variability.

As a validation approach, consistency was assessed against official planted-area series from CONAB and IBGE/PAM (state and national levels). Where available (e.g., Mato Grosso, 2023), external point/reference samples were incorporated. Accuracy reporting follows the module’s convention (overall accuracy and user/producer accuracies when independent points exist), with detailed metrics presented in the Evaluation section.

To contextualize the geographic scope, the mapping focuses on UFs where second-season corn predominates and the second season calendar is well defined (e.g., MT, MS, GO, PR, SP, MG, BA, MA, PI, TO). In contrast, Rio Grande do Sul (RS) and Santa Catarina (SC)—despite being corn producers—were not included because production there is predominantly first season, and the concept of a statewide second season is not consistently defined. These states therefore fall outside the intended scope of this product at this time.

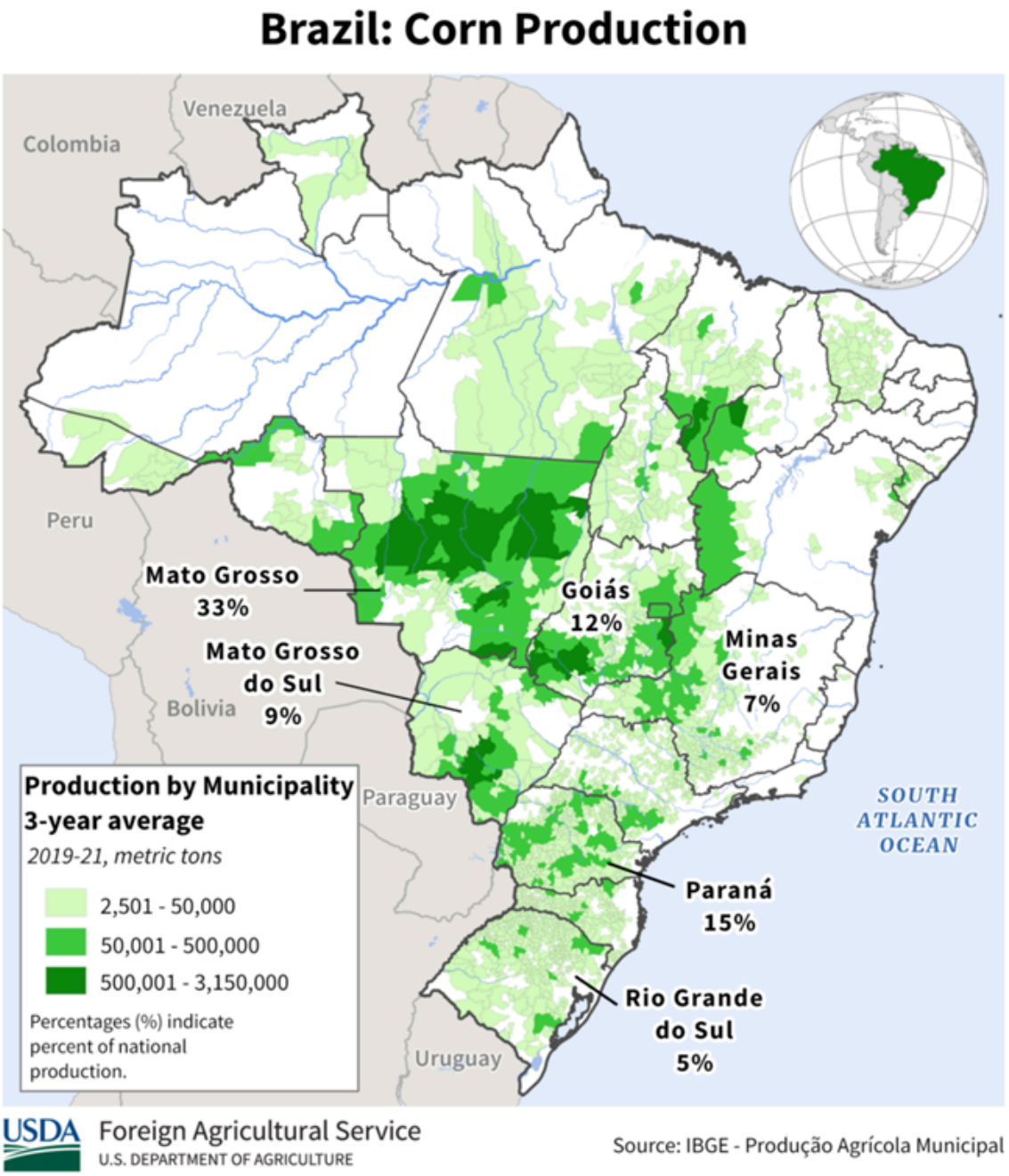


Figure 1C — Corn production in Brazil (3-year average, 2019–2021).

Municipality-level distribution of corn production highlighting concentration in central-western and southern Brazil. Figure 1C supports the selection of UFs prioritized for second season mapping and explains the exclusion of RS and SC, where corn is primarily first season.

2. Input imagery and quality assurance

Input data consist of Landsat Collection 2 surface reflectance (SR) processed over a single statewide geometry for each UF. This choice enforces uniform normalization and compositing within each state and avoids discontinuities at images' scene boundaries. For every target year, scenes are constrained to the second-season temporal window defined for that state (see Section 3), screened by CLOUD_COVER_LAND < 40%, and masked with QA_PIXEL to remove cloud, shadow, snow, and dilated cloud/water artifacts. Spectral bands are harmonized to a common set (BLUE, GREEN, RED, NIR, SWIR1, SWIR2) across sensors, and EVI2 is computed per image prior to seasonal compositing so that both reflectance and vegetation dynamics contribute to the feature space.

2.1 Sensors and allocation by period

Sensor usage is fixed by period to maximize radiometric consistency and availability within the February–May window used:

Table 1C — Sensor allocation by period (Landsat C2 SR)

Years	Sensor(s) used	Notes
2000–2002	LE07 (ETM+)	ETM+ SR; statewide geometry per UF
2003–2011	LT05 (TM)	TM SR; statewide geometry per UF
2012	LT05 + LE07	Merged collections within the UF window
2013–2025	LC08 (OLI)	OLI SR every year
2013–2025*	LC09 (OLI-2)	Included where available (e.g., 2023); 2013 = LC08 only; 2023 = LC08+LC09
*LC09 is additive to LC08 when available in the state window.		

2.2 Scene filtering and QA masks

All images within the state-specific window are first filtered by the land cloud fraction and then masked with QA bits to suppress contaminated pixels. Including the dilated cloud/water bit is important to eliminate halos adjacent to cloud/water features that otherwise propagate into the seasonal composite.

Criterion	Rule / Bit (QA_PIXEL)	Purpose
Land cloud cover	CLOUD_COVER_LAND < 40%	Coarse scene prefilter
Cloud	bit 3 = 1	Remove cloud cores

Cloud shadow	bit 4 = 1	Remove shadows
Snow	bit 5 = 1	Remove snow/ice flags
Dilated cloud / water halo	bit 2 = 1	Remove morphological halos (buffers)

Table 2C — QA rules used in masking

2.3 Band harmonization and index computation

To ensure a consistent feature space across sensors, SR bands are mapped to common names before compositing. The EVI2 index is computed per image and appended to the stack.

Sensor	Input SR bands	Common names used
LT05 / LE07	SR_B1, SR_B2, SR_B3, SR_B4, SR_B5, SR_B7	BLUE, GREEN, RED, NIR, SWIR1, SWIR2
LC08 / LC09	SR_B2, SR_B3, SR_B4, SR_B5, SR_B6, SR_B7	BLUE, GREEN, RED, NIR, SWIR1, SWIR2

Table 3C — Band mapping to common names

The per-image EVI2 layer is carried forward to seasonal reduction alongside the six reflectance bands, guaranteeing that the seasonal mosaic summarizes both spectral levels and within-window vegetation behavior.

3. Temporal windows

Temporal windows were designed to isolate the vegetative peak of second-season crops while remaining simple and reproducible across years and sensors. A baseline 1 February–31 May window was established during the pilot and then fixed per state (UF) after two checks: (i) the official agricultural calendar (to anchor sowing/harvest timing) and (ii) visual inspection of monthly Landsat mosaics (to confirm the greenness peak within the second-season period).

3.1 State-specific configuration

Final windows are specified as day–month ranges and remain unchanged across the time series. The same configuration applies to the epoch training years (2007, 2012; 2013, 2023) to maintain feature comparability with their respective prediction years.

UF	Start (DD-MM)	End (DD-MM)
BA	01/abr	30/ago
GO	01/fev	31/mai
MA	01/fev	31/mai
MG	01/fev	31/mai
MS	01/fev	30/jun
MT	01/fev	31/mai
PI	01/fev	30/jun
PR	01/fev	31/jul
SP	01/mar	30/jun
TO	01/fev	31/mai

Table 4C — Second-season windows by state (fixed for 2000–2025).

4. Sample collection and labeling

Training data were assembled to reflect second-season phenology and the spectral signature of target crops within the UF-specific window. Three classes were used throughout: (1) corn, (2) cotton, (3) other temporary crops (aggregated). Sampling and labeling were performed per state over the statewide geometry, using the same fixed window later employed in prediction.

4.1 Corn sample selection

Corn points were labeled by combining EVI2 time-series behavior with the characteristic spectral response at peak vegetative stage:

- **Phenology (EVI2 curve)**: Candidate pixels must show a well-defined rise and peak within the UF window, followed by a decline consistent with second-season corn. The curve is inspected at monthly cadence; pixels with flat or multi-modal signatures inconsistent with a single second-season cycle are excluded.
- **Spectral signature at peak (NIR–SWIR1–RED composite):** At the vegetative peak, corn fields exhibit a reddish/orange tone in false-color composites (NIR/ SWIR1 / RED), driven by high NIR reflectance (vigorous canopy), moderate SWIR and lower RED. Candidate pixels were verified visually against this tone at/near the EVI2 maximum.

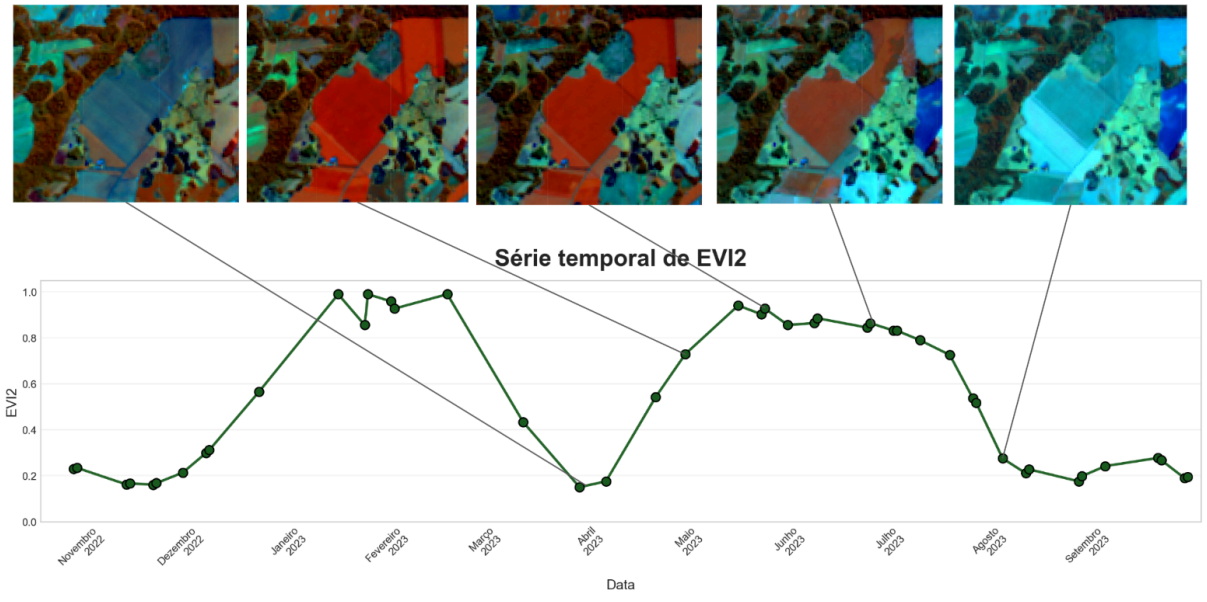


Figure 2C — Corn labeling cues. Example monthly sequence (top) and EVI2 curve (bottom) illustrating (i) the single pronounced peak within February–May and (ii) the reddish/orange response in NIR–SWIR1–RED at the peak.

4.2 Cotton and "other" samples

Cotton points were labeled with the same two-step logic—temporal consistency in EVI2 within the window and textural/tonal cues in NIR–SWIR–RED (cotton typically shows distinct brightness/texture near peak and into senescence). The "other temporaries" class aggregates annual crops present in the same window that are not corn. Including it stabilizes the decision boundary and reduces confusion with non-target annuals that may share part of the spectral/temporal space.

4.3 Epoch training sets

Two composite training sets were assembled and then applied to their respective epochs:

Epoch	Training years	Primary sensors in window	Notes on sources
2000–2012	2007, 2012	2007: LT05; 2012: LT05+LE07	Manual labeling per UF; statewide window; balance by class
2013–2025	2013, 2023	2013: LC08; 2023: LC08+LC09	Manual labeling per UF; additional MT references available (2023)

Table 5C — Training years and sources.

Where available (e.g., **MT/2023**), external references supported the curation of corn/cotton seeds. Samples were extracted from the state's seasonal mosaic at 30 m resolution using the full feature list (six bands + EVI2 × {median, p20, p80}). Sampling preserved the class labels (class ∈ {1,2,3}) and the corresponding year (one of the training years in the epoch). These per-year sample sets were then merged to form a single composite training table per epoch.

4.4 Model training and inference

Random Forest with 100 trees is used as the classifier, consuming 21 features per pixel (six harmonized SR bands plus EVI2, each summarized by median, p20, p80 within the UF window). Training follows a three-class scheme (1=corn, 2=cotton/other temporaries, 3=other) and is conducted over the statewide geometry to keep normalization and compositing consistent across each UF.

Two epoch models are adopted and transferred across their respective years: 2000–2012 trained with 2007 and 2012 samples; 2013–2025 trained with 2013 and 2023 samples. For each epoch, per-UF samples from the two training years are merged into a single table, preserving class and year; when "other temporaries" are abundant, a cap is applied to avoid dominance and maintain class balance. Training years use the same UF-specific window later used in inference, ensuring feature comparability across time.

5. Annual prediction and agricultural masking

Per year × UF, the epoch model (Section 4.4) is applied to the seasonal mosaic assembled with the UF's fixed second-season window. The output is a raster in the three-class scheme {1=corn, 2=cotton/other temporaries, 3=other}. Prediction is executed over the statewide geometry, ensuring uniform normalization and avoiding scene-edge artifacts.

Predicted rasters are then restricted to agricultural areas using the MapBiomas Collection 10 integration map. The agricultural mask is implemented by remapping target/non-target classes and masking out the remainder, as summarized below.

Source class (C10 integration)	Meaning (short)	Keep in mask	Remap
39	Annual cropland	Yes	1
41	Soybean / temporary crop	Yes	1
62	Mosaic of agriculture	Yes	1
20	Sugarcane	No	0

Table 6C — Agricultural mask (MapBiomas C10 integration).

The mask is applied after classification. Pixels mapped to 0 (e.g., sugarcane and non-agricultural classes) are removed; pixels mapped to 1 retain the predicted class {1,2,3}.

6. Post-processing filters

Post-classification filtering is applied selectively to improve spatial coherence and temporal stability. The filter routine combines class-wise spatial connectivity with year-to-year consistency rules and was executed only in five UFs where it delivered measurable gains. In other UFs, maps remain as masked predictions (no filter) to preserve genuine interannual variability.

6.1 General design

Filter operates in two class-specific stages (1=corn, 2=cotton):

Pre-connectivity: Apply connected-component filtering on binary masks (8-neighbor). The minimum patch size varies by epoch: 8 pixels for 2000–2012 and 20 pixels for 2013–2024.

Temporal consistency + post-connectivity: Enforce inclusion/keep rules by year (below), then re-apply connectivity with the same year's minPatch. The final label is mutually exclusive (class 3 "other" is not filtered).

Component	Setting / value
Neighborhood	8-connected
minPatch (2000–2012)	8 pixels
minPatch (2013–2024)	20 pixels
Output per year	class ∈ {0,1,2} (0 = masked out)

Table 7C — Filter parameters

6.2 Temporal rules by block

Rules are applied after the pre-connectivity step (8-neighbor, year-specific minPatch) and separately for each class (corn=1, cotton=2). For every year, the rule produces temporary masks per class; then post-connectivity is applied with the same year's minPatch.

Year 2000 — persistence with 2001
Keep a pixel in 2000 only if that same class is also present in 2001. After this check, apply post-connectivity.
(Example: a corn pixel in 2000 is kept only if that pixel is also corn in 2001.)

Years 2001–2022 — inclusion + keep (moderate)
Two checks run for each year y:

- Inclusion: keep the pixel in y if it is mapped as that class in y, or if it appears in both $y-1$ and $y+1$. Also require that the other class is not present in y (to avoid immediate conflicts).
  (Example: for y=2010, a corn pixel is included if it's corn in 2010, or if it's corn in 2009 and 2011; and it must not be cotton in 2010.)
- Keep: once included, keep the pixel in y if there is support in any of the neighboring years $y-1$, $y+1$, $y+2$, or $y+3$ for the same class, or if there is paired support of the other class in $y-1$ and $y+1$ (this paired support term stabilizes boundaries where corn/cotton alternate across years).
  (Example: still for y=2010, corn remains if corn appears in 2009, 2011, 2012, or 2013; it also remains if cotton appears in both 2009 and 2011, which helps avoid flipping at field boundaries.)

After inclusion and keep, apply post-connectivity and enforce class precedence (cotton over corn).

Year 2023 — ±1 inclusion, with look-ahead for keep

- Inclusion: keep the pixel in 2023 if it is mapped in 2023, or if it appears in both 2022 and 2024 for the same class; also require that the other class is not present in 2023.
  (Example: a corn pixel is included if corn is present in 2023, or if corn is present in 2022 and 2024; and it must not be cotton in 2023.)
- Keep: once included, keep the pixel in 2023 if it has support in 2022, 2024, or 2025 for the same class, or paired support of the other class in 2022 and 2024.
  (Example: corn in 2023 is retained if corn appears in any of 2022, 2024, or 2025; it is also retained if cotton appears in both 2022 and 2024.)

Then apply post-connectivity and class precedence.

Year 2024 — bridge with 2023 & 2025

- Inclusion: keep the pixel in 2024 if it is mapped in 2024, or if it appears in both 2023 and 2025 for the same class; also require that the other class is not present in 2024. (Example: a corn pixel is included if corn is present in 2024, or if corn is present in both 2023 and 2025; and it must not be cotton in 2024.)
- Keep: once included, keep the pixel in 2024 if it has support in 2023 or 2025 for the same class, or paired support of the other class in 2023 and 2025. (Example: corn in 2024 is retained if corn appears in 2023 or 2025; it is also retained if cotton appears in both 2023 and 2025.)

Then apply post-connectivity and class precedence.

Year 2025 — connectivity only (no temporal blending)
For 2025, do not use any temporal rule. Keep pixels of each class that pass connectivity in 2025 and then enforce class precedence.
(Example: a corn pixel is kept in 2025 solely based on the 2025 connectivity threshold; 2024 is not consulted.)

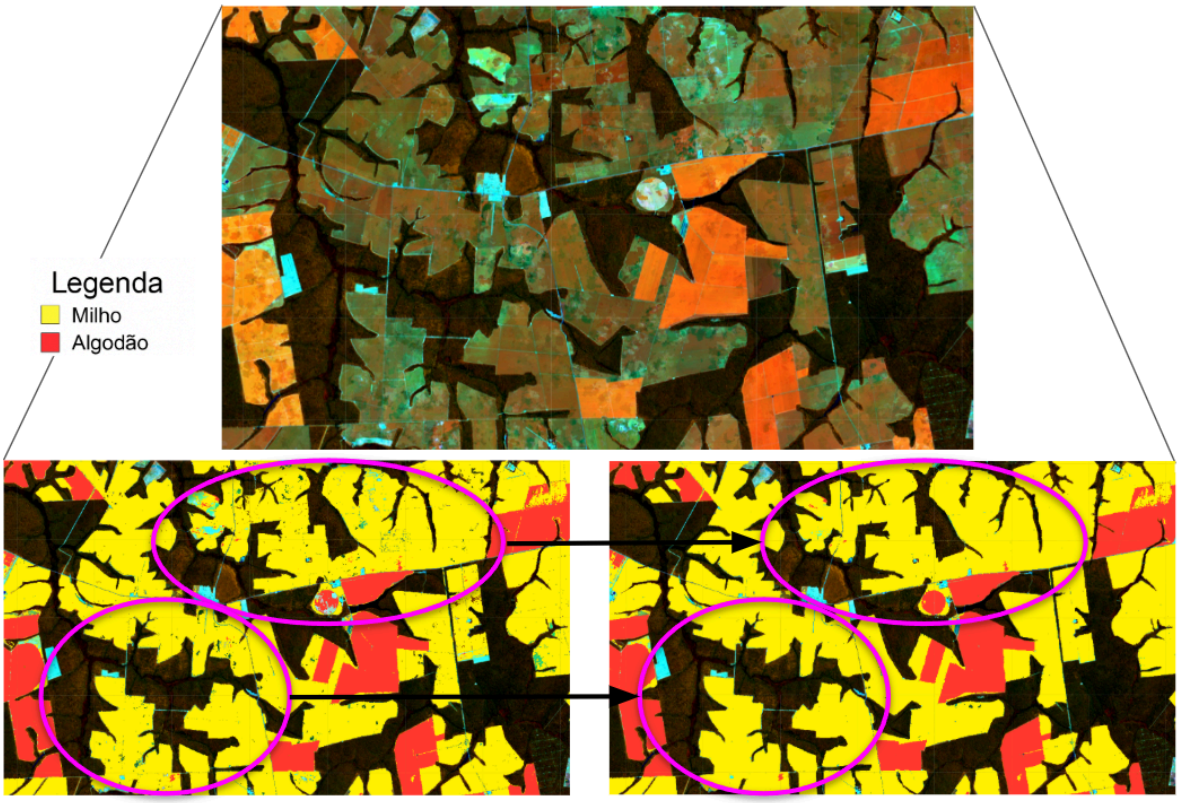


Figure 3C — Effect of the post-processing filter. Top: NIR–SWIR1–RED false-color mosaic over the study area at the second-season peak. Bottom-left: classification without filter (legend: yellow = corn, red = cotton). Bottom-right: classification with filters (8-connected components; minPatch = 20 for 2013–2025, 8 for 2000–2012). Magenta circles highlight reductions of speckle and consolidation of field patches after filtering, with improved boundary coherence and removal of small transients.

6.3 Spatial Filtering

After the multi-year temporal consistency rules described in Section 6.2 are applied, a final post-processing chain is performed on the classified second-season maps to improve spatial consistency. First, a morphological dilation operation is applied to the temporally filtered raster to fill small internal gaps and unmapped patches within the identified agricultural fields. Subsequently, to remove remaining high-frequency noise and isolated patches that do not represent consolidated field-crop areas, a minimum connectivity filter is applied. This filter removes isolated patches smaller than six pixels using an eight-neighbor connectivity configuration.

6.4 Impact and limitations

Filters reduces speckle and transient 1–2-pixel patches, improves field-level coherence, and stabilizes interannual dynamics while respecting the class hierarchy. A known trade-off is the potential erosion of very small fields (below the minPatch threshold).

7. Validation and evaluation

Validation focuses on (i) a point-based check where independent labels exist and (ii) temporal consistency against official statistics. All assessments use the final product (after agricultural mask and, where applicable, filters).

Independent samples are available only for Mato Grosso (MT), 2023 (internal set). These homogeneous reference polygons (100% corn or 100% cotton) were used in a pixel-based validation approach, where every 30m pixel within validation polygons was treated as an independent sample. This provides a more rigorous assessment of classification performance at the native resolution of the product.

7.1 Spatial accuracy assessment

The validation dataset comprised over 6.3 million pixels across 2,331 homogeneous reference polygons in Mato Grosso for the 2023 season. Both the original classification and filtered product were evaluated using pixel-based accuracy assessment.

Metric	Original Classification	Filtered Product
Overall Accuracy	98.26%	98.65%
Corn Producer's Accuracy	98.3%	98.7%
Corn User's Accuracy	99.97%	99.98%
Cotton Producer's Accuracy	97.5%	98.1%
Cotton User's Accuracy	43.0%	48.1%
Total Pixels	5,734,439	6,305,627

Table 8C- Pixel-based accuracy metrics for corn and cotton classification in Mato Grosso, 2023

	Reference: Corn	Reference: Cotton	User's Accuracy
Predicted: Corn	5,561,256	1,859	99.97%
Predicted: Cotton	97,69	73,634	43.0%
Producer's Accuracy	98.3%	97.5%	

Table 9C - Original Classification Confusion Matrix

	Reference: Corn	Reference: Cotton	User's Accuracy
Predicted: Corn	6,142,745	1,535	99.98%

Predicted: Cotton	83,686	77,661	48.1%
Producer's Accuracy	98.7%	98.1%	

Table 10C - Filtered Product Confusion Matrix

The pixel-based validation reveals excellent detection capability for both crops, with producer's accuracy exceeding 98% for corn and cotton in the filtered product. However, significant overestimation of cotton is evident, with user's accuracy around 48%. The filtering process improved overall accuracy by 0.4 percentage points and substantially enhanced cotton user's accuracy from 43% to 48%, demonstrating the value of the applied post-processing steps.

References

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Atlas irrigação: uso da água na agricultura irrigada /Agência Nacional de Águas. Brasília: ANA, 2017.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Levantamento da agricultura irrigada por pivôs centrais no Brasil/Agência Nacional de Águas, Embrapa Milho e Sorgo. 2. ed. Brasília: ANA, 2019a.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Polos nacionais de agricultura irrigada: mapeamento de áreas irrigadas com imagens de satélite/Agência Nacional agricultura irrigada por pivôs centrais de irrigação no Brasil - 1985 - 2022/ Agência Nacional de Águas - Brasília: ANA, 2022. Boletim do SNIRH n. 4.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Mapeamento do arroz irrigado no Brasil. Brasília: ANA & Conab, 2020.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Atlas irrigação: uso da água na agricultura irrigada (2ª edição)/Agência Nacional de Águas - Brasília: ANA, 2021a.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Mapeamento do arroz irrigado no Brasil/Agência Nacional de Águas, Companhia Nacional de Abastecimento. Brasília: ANA, 2021b.

AGÊNCIA NACIONAL DE ÁGUAS (ANA). Levantamento da Agricultura Irrigada por Pivôs Centrais no Brasil. Boletim do SNIRH 4. 2023.

CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. Mapeamento Agrícolas. Available at: <https://www.conab.gov.br/info-agro/safras/mapeamentos-agricolas?start=20>.

EMPRESA BRASILEIRA DE PESQUISA E AGROPECUÁRIA (EMBRAPA). Dados conjunturais da produção de arroz (Oryza sativa L.) no Brasil (1986 a 2019): área, produção e rendimento. Santo Antônio de Goiás: Embrapa Arroz e Feijão, 2020. Available at: <http://www.cnpaf.embrapa.br/socioeconomia/index.htm>. Access on : 07/04/2022.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). 2009. Censo agropecuário 2006: resultados definitivos. Rio de Janeiro: IBGE, 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). 2019. Censo agropecuário 2017: resultados definitivos. Rio de Janeiro: IBGE, 2019.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., ... & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific data*, *2*(1), 1-21.

Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, *58*(3), 257-266.

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaudoise Sci Nat*, *37*, 547-579.

Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote sensing of Environment*, *112*(10), 3833-3845.

Jönsson, P., & Eklundh, L. (2004). TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & geosciences*, *30*(8), 833-845.

Nagler, P. L., Inoue, Y., Glenn, E. P., Russ, A. L., & Daughtry, C. S. T. (2003). Cellulose absorption index (CAI) to quantify mixed soil–plant litter scenes. Remote Sensing of Environment, 87(2-3), 310-325.

Ronneberger, O.; Fischer, P.; Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention; Springer: Berlin, Germany, 2015; pp. 234–241.

Rouse, J., Haas, R., Schell, J., & Deering, D. (1974). Monitoring vegetation systems in the great plains with ERTS. In *Proceedings of the Third Earth Resources Technology Satellite—1 Symposium; NASA SP-351* (pp. 309-317).

Saraiva, M., Protas, É., Salgado, M., & Souza Jr, C. (2020). Automatic Mapping of Center Pivot Irrigation Systems from Satellite Images Using Deep Learning. Remote Sensing, 12(3), 558.

Whittaker, E. T. (1922). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, *41*, 63-75.

Xu, H. (2006). Modification of normalized difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, *27*(14), 3025-3033.

Zhou, J., M. Menenti, L. Jia, B. Gao, F. Zhao, Y. Cui, X. Xiong, X. Liu, and D. Li. A Scalable Software Package for Time Series Reconstruction of Remote Sensing Datasets on the Google Earth Engine Platform. International Journal of Digital Earth, v. 16(1), pp. 988–1007, 2023